Collective Decisions and Voting

Nicolaus Tideman

COLLECTIVE DECISIONS AND VOTING

This page intentionally left blank

Collective Decisions and Voting

The Potential for Public Choice

NICOLAUS TIDEMAN
Virginia Polytechnic Institute and State University, USA

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2006 by Ashgate Publishing

Published 2016 by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN
711 Third Avenue, New York, NY 10017, USA

Routledge is an imprint of the Taylor & Francis Group, an informa business

Copyright © 2006 Nicolaus Tideman

Nicolaus Tideman has asserted his right under the Copyright, Designs and Patents Act, 1988, to be identified as the author of this work.

Permission for the extensive quotations in Chapter 10 from Kenneth Arrow's *Social Choice and Individual Values* (© 1963) has been granted by Yale University Press.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

Notice:
Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

British Library Cataloguing in Publication Data
Tideman, Nicolaus
Collective decisions and voting : the potential for public choice
1.Social choice 2.Elections 3.Plebiscite 4.Group decision making
I.Title
302.1'3

Library of Congress Cataloging-in-Publication Data
Tideman, Nicolaus.
Collective decisions and voting : the potential for public choice / by Nicolaus Tideman.
p. cm.
Includes bibliographical references and index.
ISBN-13: 978-0-7546-4717-1
ISBN-10: 0-7546-4717-X
1. Voting. 2. Group decision making. 3. Social choice. I. Title.

JF1001.T49 2006
324.6--dc22

2006021134

Transfered to Digital Printing in 2012

ISBN 9780754647171 (hbk)

Contents

List of Figures *vii*
List of Tables *ix*
Preface *xi*
Acknowledgements *xvii*
List of Abbreviations *xviii*

PART 1 Collective Decisions

1 Defining a Collective Decision 3
2 A Taxonomy of Collective Decision Procedures 9
3 Economic Criteria for Evaluating Collective Decisions 23
4 General Criteria for Evaluating Collective Decision Procedures 35
5 Relative Advantages of Modes of Making Collective Decisions 43
6 What is a Good Collective Decision? 57

PART 2 Voting

7 An Overview of Voting 77
8 Majority Rule and its Weighted Analog 83
9 Voting Cycles 93
10 The Arrow Theorem 123
11 Strategic Voting and the Gibbard–Satterthwaite Theorem 143
12 Criteria for Evaluating Ranking-Based Vote-Processing Rules 149
13 Vote Processing Rules for Selecting One Option from Many When Votes Have Predetermined Weights: Alternatives to Plurality 165
14 Vote Processing Rules for Selecting One Option from a Continuum of One or More Dimensions When Votes Have Predetermined Weights 245

15 Vote Processing Rules for Selecting More than One Candidate When Votes Have Predetermined Weights: Proportional Representation 263

16 Vote Processing Rules With Endogenous Weights for Self-Interested Advocates: Revealing Intensities of Preferences 289

17 Lessons from the Excursion 321

Bibliography *327*
Index *333*

List of Figures

2.1	Locations of modes in the agreement-on-outcome family	14
2.2	Locations of sub-modes in the agreement-on-outcome family	15
9.1	An example of single-peaked preferences	96
9.2	Options in a plane, preferences determined by distance	98
9.3	Frequency of elections containing at least one cyclic anomaly as a function of the number of candidates in the election	103
9.4	Frequency of cyclic anomalies in components of matrices of majorities as function of the number of voters who compare a pair of candidates	104
9.5	Distribution of majorities for components of matrices of majorities representing cyclic anomalies	105
9.6	Frequency of cyclic anomalies in components of matrices of majorities as a function of relative distance from the main diagonal	105
9.7	Frequency of cyclic anomalies in matrices of majorities as a function of relative distance from the center of a diagonal	106
9.8	An example of the beta density function, with $r = 0.3$ and $n = 50$	109
13.1	Scatter plot of the percentages of voters needed and available for minimal strategies under plurality, for 87 elections	167
13.2	Scatter plot of the percentages of voters needed and available for minimal strategies under approval voting, for 87 elections	173
13.3	Scatter plot of the percentages of voters needed and available for minimal strategies under range voting, for 87 elections	176
13.4	Scatter plot of the percentages of voters needed and available for minimal strategies under the Borda rule, for 87 elections	180
13.5	Scatter plot of the percentages of voters needed and available for minimal strategies under the Condorcet rule, for 87 elections	189
13.6	Scatter plot of the percentages of voters needed and available for minimal strategies under two-ballot majority rule, for 87 elections	193

13.7 Scatter plot of the percentages of voters needed and available for minimal strategies under the alternative vote, for 87 elections 195

13.8 Scatter plot of the percentages of voters needed and available for minimal strategies under Bucklin voting, for 87 elections 206

13.9 Scatter plot of the percentages of voters needed and available for minimal strategies under the Copeland rule, for 87 elections 209

13.10 Scatter plot of the percentages of voters needed and available for minimal strategies under the Black rule, for 87 elections 211

14.1 Strategic opportunities under the UMC rule 246

14.2 Non-independent preferences in two dimensions 249

14.3 Outcomes that vary with rotations of axes under the MMC rule 251

14.4 An illustration of the MBP rule 252

14.5 The incentive to misstate preferences under the MSD rule 256

16.1 Probabilistic prospects for three options 303

16.2 Valuations of movements among three lattice points 314

List of Tables

2.1	A taxonomy of collective decision procedures	9
4.1	A taxonomy of criteria for evaluating collective decision procedures	35
5.1	The attractiveness of modes in terms of criteria	44
7.1	The dimensions of voting situations	80
9.1	Values of $f(p)$ and $g(p)$ for the example in the text	96
9.2	A summary of characteristics of 87 elections employed in investigating the probability of a voting cycle	101
9.3	Average proportion of candidates ranked, sorted by number of candidates and number elected (all elections weighted equally)	101
9.4	An example of a matrix of majorities	102
9.5	Average value of z as a function of location in the matrix of majorities	107
9.6	Implied value of n as a function of location in the matrix of majorities	111
9.7	Estimated parameters of the determinants of the probability of a cycle	113
9.8	r as a function of i and j for the case of 10 candidates	114
9.9	Estimated probability of a tie between the top two candidates	116
9.10	Estimated probability of some tie between adjacent candidates	117
9.11	Estimated probability of a majority-rule equilibrium	118
9.12	Estimated probability of a dominant candidate	119
9.13	Estimated probability of no cycles anywhere	120
9.14	Estimated probability of no anomalies anywhere	121
12.1	A classification of previously encountered properties of RBVPRs by frequency of occurrence in potentially interesting rules	149
12.2	A taxonomy of properties used to evaluate RBVPRs	151
13.1	Properties of vote-processing rules	237

13.2 A classification of vote-processing rules by supportability 238

14.1 Properties of six rules for selecting one option from a continuum 261

14.2 The choice of a vote-processing rule for a continuum of two or more dimensions as a function of two criteria 261

15.1 One row of a CPO-STV matrix of majorities 281

Preface

How do people coordinate their intentions? How can the mechanisms that they use be classified? Are there other mechanisms that would be attractive if people knew about them? And what about the moral dimension of collective decisions? How is it possible to know what a group should do when the members of the group disagree among themselves? When, if ever, is it possible to justify a claim that individuals ought to abide by collective decisions with which they disagree? These are questions that are addressed in this inquiry into collective decisions and voting.

At one level, this book is a smorgasbord of suggestions about voting procedures and other ways of making collective decisions, suggestions that might be of interest to people who would like to change collective decision procedures in ways that would potentially be satisfying to all concerned. It provides a set of definitions, taxonomies, properties, advantages and disadvantages of a wide variety of procedures. At another level, this book is simply a report on puzzles with respect to voting that have intrigued me, and the resolutions that I have found for them. The subsequent chapters of the book develop these ideas in a logical sequence. Here I offer a more historical account of my interest in them, resulting in a sequence that is roughly backward from the sequence of chapters in the book.

My earliest recollection of an interest in voting goes back to a day when I was eight years old, in San Francisco. Looking at the election results published in the San Francisco Chronicle, I noted that nine persons had been elected to the Board of Supervisors. The method of election was that each voter was allowed to vote for four candidates (or was it five?), and the nine with the most votes were the winners. I asked my father why they did it that way. He couldn't give me a satisfying answer, and the question of how a group of persons ought to be elected stuck with me. Eventually I learned that some people thought that the Single Transferable Vote was a satisfying way of electing a group. I could see elements of merit in that procedure, but there were obvious problems as well. In the mid 1980s I spent many dozens of hours on the question of what would be a satisfying way to elect a group, and finally came up with a recipe for comparing pairs of outcomes by the Single Transferable Vote,[1] which is discussed in Chapter 15, in the context of various ways of electing a group and why one way or another might be most satisfying.

When I was 12 years old I was nominated to be treasurer of my class at school. Being a whiz a math, I thought I would be a fine treasurer and salivated for the position. Michelle was also nominated. I don't remember whether I nominated

Michelle's best friend Charlotte, or whether I seconded the nomination. But I did one or the other, and my reaction was instinctive. I knew what I was doing. In the election, Charlotte received 11 votes, Michelle received 12, and I received 13 and became treasurer. I was immediately aware that I had won the office in a way that was not objectively fair and reasonable. The thought passed through my head that the honorable thing to do would be to suggest a run-off between Michelle and me. But I wanted the office too badly. Furthermore, there is some question as to whether such an innovation in the rules is proper in the middle of an election procedure. So there was no run-off. But the event imprinted on me the problem of what procedure should be used to determine the winner of an election when there are more than two candidates. In the early 1980s, when my Department was seeking to aggregate the opinions of its members as to whom it should hire, I found that the procedure that seemed most sensible to me was the "ranked pairs" rule,[2] which is discussed in Chapter 13, in the context of a long list of rules for electing one candidate, their properties, advantages and disadvantages. Chapter 12 provides definitions and explanations of the properties and criteria that are employed in Chapter 13.

As an undergraduate, I learned that there was something called welfare economics, which yielded conclusions that particular economic outcomes were good. That astounded me. It suggested that when people disagreed about what ought to be done, it would be possible for one group to say to another, "We have proven that what we want is good and what you want is not. Don't give us a hard time." I could not imagine how such an assertion could be valid. I wanted to understand how the rabbit got into the hat. Eventually I learned that if one posits normative axioms, then it is possible to use accepted logic to derive normative conclusions. A proof is simply a correspondence between assumptions and conclusions developed o the basis of accepted principles of logic. It establishes that something is true only if one knows that the axioms are true. And securing agreement that normative axioms are true is very difficult. I learned that there were at least four distinct normative traditions in twentieth century economics. These are reviewed in Chapter 3. My own view of what constitutes an adequate basis for disappointing someone who desires a particular outcome is developed in Chapter 6.

As an undergraduate, I also learned that there was something called the Arrow Theorem that showed that democracy was inevitably incoherent, or some such thing. It seemed impossible that I would ever be able to comprehend such a theorem, so while I was curious about it, I left it unexplored. There were further occasional references to it in my graduate education, but no explicit treatment of it. So I was quite startled when, as a new faculty member at Harvard, one of the first things I was expected to do was to explain the Arrow theorem to students in a course in intermediate microeconomic theory. Fortunately, there were graduate students who were ready to fill me in on what I needed to know. Unfortunately, I could not comprehend the theorem to my satisfaction and was quite dissatisfied

with the explanation I gave to the students in that class. Over the years, I continued to puzzle over what was actually going on in the Arrow theorem. I looked long at hard at Arrow's proof, as well as at proofs of similar theorems. I developed an idea about what was going on in the theorem, but I was aware that my ideas were not congruous with the general consensus of economic theorists. In the late 1980s, deciding that I had to resolve the matter, I wrote letters to a number of economic theorists requesting that they engage me in dialogue on the question of what the Arrow theorem meant. One person to whom I wrote, Georges Bordes, replied with a 12-page hand-written letter. I replied with a letter half as long, and he replied with another long letter. We each learned something from the other, so we decided to collaborate on a joint paper on what the Arrow Theorem really means.[3] The understandings that I thereby acquired regarding of the meaning of the Arrow Theorem and its significance for the limits of a coherent democracy are explained in Chapter 10, in the context of a detailed parsing of Arrow's proof.

The Arrow Theorem is only half of the story of the logical limits of democracy. The other half is the Gibbard–Satterthwaite Theorem. This theorem employs the Arrow Theorem to prove that no voting procedure for more than two options can motivate voters in all circumstances to report their rankings of the options truthfully. As soon as I heard about this theorem, which was first published in the mid 1970s, I realized that it was very important, but I did not expect to be able to understand the proof. The first few times I presented it to students, I understood each step, but not its overall structure. Eventually I came to understand not only the structure of the proof, but also why one should expect that the theorem would be true. I have sought to convey that understanding in the presentation and discussion of the Gibbard–Satterthwaite Theorem in Chapter 11.

The Arrow Theorem and the Gibbard–Satterthwaite Theorem are both closely tied to the phenomenon of majority-rule cycles. Thus the significance of these theorems is related to the question of how frequently cycles can reasonably be expected to occur in real voting situations. It turns out that cycles will be common if not pervasive if there are no restrictions on placing redistributive proposals on the agenda, or if there are relatively few voters or relatively many options. But in typical voting situations with relatively few options and more than several hundred voters, cycles will be quite rare and can be dealt with in a way similar to the way that ties are dealt with. This is explained in Chapter 9.

One circumstance in which cycles cannot arise is if the options under consideration lie in a one-dimensional continuum and voters have "single-peaked" preferences in that dimension. In that case there is a majority-rule equilibrium at the preference of the median voter. There is in general no majority-rule equilibrium when the options lie in a continuum of more than one dimension, but there are nevertheless some voting procedures that might be attractive when the options lie in a continuum of several dimensions. These are explored in Chapter 14.

Discussions of voting usually presume without discussion that all votes will be weighted equally. (The one exception is voting on matters affecting a corporation, where the general presumption is that all shares will be weighted equally.) In my doctoral dissertation I dealt with an issue (proposals for changes in land use) where it is arguably sensible to weight votes according to prior estimates of the relative intensities of the voters' preferences.[4] The general question of when it might be attractive to employ weighted voting, and why, has intrigued me over the years. In Chapter 8 I explore this question, along with the question of where ordinary majority rule for choices between two options came from, and what rationales it has.[5]

While the suggestion that I made in my doctoral dissertation was for weighting according to estimates of the intensities of voters' preferences, what I would have suggested if I could have imagined it was feasible was weighting according to the actual intensities of voters' preferences. But I could not imagine that such a thing was feasible. I could not imagine it despite the fact that one of my fellow graduate students, Edward Clarke, informed me, as I worked on my dissertation in 1968, that he had found a mechanism that motivated voters to report honestly the intensities of their preferences for public goods. I was sure he was mistaken and declined to take the time to examine his claim. But Clarke's ideas kept crossing my path.

In the 1970-71 academic year I was invited to contribute a chapter to the book *Public Prices for Public Products*. I contributed a chapter on a mechanism for successive revisions in the level of a public good.[6] Clarke offered a version of his idea in the following chapter.[7] That same year my dissertation advisor, George Tolley, who was also Clarke's advisor, invited Clarke and me to give back-to-back presentations of our ideas in a seminar. I had no understanding of Clarke's ideas after that seminar.

In the fall of 1973 I embarked on a post-doctoral fellowship at the Center for the Study of Public Choice, at Virginia Polytechnic Institute and State University. I told Gordon Tullock, who was also at the Center, that I planned to spend my post-doctoral year reviewing ways of making collective decisions. His response was, "Well, if you are going to do that, you really ought to include Ed Clarke's ideas in your review." He said that he had published Clarke's idea in Public Choice[8] because it was important if it was right, but he did not understand it. (Tullock made most editorial decisions without consulting referees.) I reluctantly agreed to include it.

Three months into my post-doctoral fellowship, I agreed to teach and postpone the postdoctoral fellowship to the following year, because of an unexpected shortage of teachers, so it was not until the spring of 1975 that I confronted the need to deal with Clarke's ideas. And I still might have let it slip by if I had not encountered Martin Bailey at the meetings of the Public Choice Society that spring. I told Bailey what I was doing, and he told me that Clarke was right. With

that endorsement and the knowledge that I had said I would deal with it, I sat down to determine what was going on in Clarke's papers.

After ten minutes or so of concentrated effort, things suddenly clicked. I saw that Clarke had indeed devised a way of motivating people to report honestly the intensities of their preferences for public goods. I quickly organized an impromptu seminar to explain the ideas to others. Five minutes into my presentation, Gordon Tullock started finishing my sentences. He said that the idea needed a new exposition, so he and I agreed to collaborate on that. We named the idea the "demand revealing process".[9] And that is how I happened to write my least original paper, which also became my most widely cited paper. Over the years, my understanding of the demand revealing process has evolved. My current understanding is presented in Chapter 16.

So far I have described a variety of ideas about voting with only occasional connections among them. Once I had all of the ideas in hand, it was natural to look for other ways in which they were connected. The result was a description of four different dimensions in which voting situations could differ. Each of the ideas that had captured my interest could be identified with a location in this four-dimensional space. This is explained in Chapter 7.

When I first conceived the idea of this book, I realized that voting was not the only way to make collective decision, and that a thorough account of voting should include a discussion of the relationship between voting and other ways of making collective decisions. After pondering alternative classifications of ways of making collective decisions for a number of years, I realized that there was a natural two-level taxonomy. This is explained in Chapter 2.

As soon as one sees all the different ways in which collective decisions might be made, the question that naturally arises is, "Why do we use the procedures we use instead of the available alternatives?" But before this question can be addressed, one must know what a good outcome of a collective decision procedure is. Thus Chapter 4 reviews criteria for evaluating collective decision procedures, and Chapter 5 applies these criteria to the ways of making collective decisions discussed in Chapter 2.

All the concern for collective decisions caused me to think about what a collective decision really is. So in Chapter 1 I define a decision, a collectivity, and hence a collective decision. The reasoning may surprise you.

Nicolaus Tideman

Notes

1 Nicolaus Tideman, "The Single Transferable Vote," *Journal of Economic Perspectives* 9/1 (1995): 27–38.

2 T.N. Tideman, "Independence of Clones as a Criterion for Voting Rules," *Social Choice and Welfare* 4 (1987): 85–206.
3 Georges Bordes and Nicolaus Tideman, "Independence of Irrelevant Alternatives in the Theory of Voting," *Theory and Decision* 30 (1991): 63–86.
4 T. Nicolaus Tideman, "Land Use Control through Administered Compensation," *Proceedings of the Meetings of the American Real Estate and Urban Economics Association* (1969): 287–93.
5 Nicolaus Tideman, "A Majority-Rule Characterization with Multiple Extensions," *Social Choice and Welfare* 3 (1986): 17–30.
6 Nicolaus Tideman, "Efficient Provision of Public Goods," in S. Mushkin (ed.), *Public Prices for Public Products*, (Washington, 1972), pp. 111–23.
7 Edward H. Clarke, "Multipart Pricing of Public Goods: An Example," in S. Mushkin, (ed.), *Public Prices for Public Products*, (Washington, 1972), pp. 125–30.
8 Edward H. Clarke, "Multipart pricing of Public Goods," *Public Choice* 11, (1971): 17–33.
9 Nicolaus Tideman and Gordon Tullock, "A New and Superior Process for Making Social Choices." *Journal of Political Economy* 84 (1976), 1145–59.

Acknowledgements

A number of people provided significant help with this book. When Georgia Pourchot served as my research assistant she read drafts diligently, pointed out weak places diplomatically, and kept after me to identify the next thing that needed to be done and do it. Ed Clarke gave me repeated gentle encouragement. James Green-Armytage provided detailed comments on a draft and helped connect me to the on-line discussion of voting. Markus Schulze helped ensure that I had a proper understanding of the voting system that bears his name. Finally, I acknowledge the loving support of my wife, Estill Putney.

List of Abbreviations

Alt.	Alternative
CPO-STV	comparisons of pairs of outcomes by the single transferable vote
ERS	Electoral Reform Society of Great Britain and Ireland
Est.	estimated
GOCHA	Generalized Optimal-CHoice Axiom
IIA	independence of irrelevant alternatives
Invar.	Invariant
Maj.	Majority
MBP	multidimensional balanced pulls
MMBP	multistage multidimensional balanced pulls
MMC	multidimensional median choice
MMSF	multistage minimum sum of forces
MSD	minimum sum of distances
MSF	minimum sum of forces
OWS	one-way streets
PSC	proportionality for solid coalitions
RBVPR	ranking-based vote-processing rule
Simp.	Simplified
SQ	status quo
STV	single transferable vote
TBRO	tie-breaking ranking of the options
TBRP	tie-breaking ranking of pairs
TL	traffic lights
UD	universal domain
UMC	unidimensional median choice
Wt.	Weighted

PART 1
Collective Decisions

This page intentionally left blank

Chapter 1

Defining a Collective Decision

The Importance of Choosing the Right Definition

In *Through the Looking Glass*, Humpty Dumpty says that when he uses a word, it means exactly what he chooses it to mean, neither more nor less.[1] Our reactions to his peculiar definitions reveal that a writer who wishes to communicate cannot claim Humpty Dumpty's freedom. Readers generally have some previous acquaintance with the terms that are to be defined, and a definition that takes no account of this leaves them alienated or confused.

A good definition builds upon and sharpens distinctions with which readers are already familiar. Ideally, a definition is not imposed by the author, but rather greeted by readers with the thought, "Yes, that is exactly how the term ought to be understood."

This chapter advances a definition of a collective decision, a term that suggests meanings to most readers even if they are not prepared to define it. My task is to persuade you that the understandings you already have of this term are best captured in a definition that may be surprising. But the definition of a collective decision will seem natural after the definitions of "decision" and "collectivity," which are developed in the next two sections.

Defining "Decision"

A decision is not a physical event, but rather a mental one. If you say, "I have decided what to do," it does not imply that you have actually done anything, only that your mind is made up. Thus decisions are not observed directly; rather, they are inferred from speech or other actions. For this reason, there is an unavoidable element of uncertainty in any assertion that a decision has been made.

From its Latin root, to "decide" is to "cut away." The applicability of deciding as cutting away can be seen when the future is conceived as a branching tree of possibilities. The actual course that events will take depends on a combination of one's own choices, the choices of others, and those of "nature." A decision cuts away some of the branches.

The actions involved in implementing decisions need not be physical. While flying in an airplane I may decide to pray, or to compute in my head the number of square feet in an acre, and the implementation of my decision might be

unobservable to anyone watching me. Therefore the word "action" will be stretched to cover all behavior that implements decisions, whether the behavior is physical or mental.

Sometimes a decision involves the selection of a single intended course of action ("Tomorrow I will go to the city"), contingent on circumstances, such as continued good health, that are often left unspecified. At other times a decision represents the exclusion of a possibility ("I will not go to the city tomorrow") or the narrowing down of possibilities ("For my vacation I will go either to Boston or to New York"). All are examples of pruning the tree of future possibilities.

Even if there is an interval of time between the formation of an intention and the corresponding action, there is still no decision if the actor does not see himself or herself as making a choice. A soldier following orders or a driver responding to a traffic signal may act on the basis of a habit of behaving as directed in these circumstances, rather than as a result of a decision to do so on this occasion, although he might also think about it and decide to follow the orders or the traffic signal.

These considerations lead to the following definition of a decision.

> *Definition*: Decision: A process that focuses intention, contingent frequently on conditions that are not fully specified, on a restricted set of future choice options, often on just one choice option.

With "decision" defined, it is now time to define "collectivity."

Defining "Collectivity"

A collectivity is a group of individuals who recognize the possibility of coordinating their actions usefully. A team, a club, a gang of thieves, a union, a church, a family, the citizens of a polity, and the participants in an exchange are all examples of collectivities.

The members of a collectivity need not be human, unless, which seems unlikely, only humans possess consciousness. When a dog picks up a stick and drops it at one's feet, the dog is saying, in effect, "I would enjoy fetching the stick if you would be willing to throw it." The dog is seeking to establish a collectivity. The language ability of chimpanzees has been studied by putting pairs of chimpanzees in situations where they are mutually rewarded when one passes information to the other through symbols.[2] The chimps' ability to succeed in such situations strongly suggests that they possess an explicit recognition of the possibility of coordinating their actions usefully. A flock of birds, on the other hand, probably is not a collectivity. They probably lack both perceptions of themselves as separate beings and explicit perceptions of the benefits that might be obtained by coordinating their actions.

The relations among the members of a collectivity need not be symmetric (they can have quite disparate roles), nor do the members need to enjoy the prospect of relating to one another. A mugger and his victim can form a collectivity. If the mugger would prefer not to shoot the victim, but would do so if that were necessary to obtain the victim's money, and if the victim would rather part with his money than be shot, then there is a potential for coordinated action between mugger and victim that would benefit both, by comparison with the possible results of uncoordinated action. In these circumstances the mugger and his victim form a collectivity. On the other hand, a person who snatches a purse and runs away does not form a collectivity with his victim, because there is no potential for a mutual desire to coordinate their actions.

The following definition captures what has been said about collectivities.

> *Definition*: Collectivity: a set of conscious individuals who perceive themselves as separate beings and also perceive benefits that might be obtained by coordination among themselves.

With "decision" and "collectivity" defined, it is now reasonably straightforward to define a collective decision.

Defining "Collective Decision"

A collective decision occurs when members of a collectivity make individual decisions that they would not make if the other members of the collectivity were not making related decisions. A collective decision thus entails coordination of intentions.

The possibility of uncertainty or ambiguity in communication raises difficulties for the definition of a collective decision. It is certainly not necessary for all of the members of the collectivity to intend to act in the manner prescribed by a presumed collective decision in order for a collective decision to be made. For example, a club may decide to have a picnic on a certain day, and one of the members of the club may decide not to participate, without making this known, and still it would be appropriate to say that there had been a collective decision to have a picnic. It could be said that the collectivity that decided to have the picnic was not the full club, but only the members who intended to go. However, if so few members decided to go that it was not possible to describe the club as having had a picnic, then we would say that what might have appeared to be a collective decision, in fact was not one.

Ambiguity in whether a collective decision has been made can lead to disputes in other contexts. Consider the following dialogue:

"My car has been giving me so much trouble. It's totally worthless."

"I'll give you $10 for it."

"Sold."

If the "buyer" shows up with $10 and wants the car, is the "seller" obliged to sell, or can he say he was just joking and keep the car? What if the amount of money was $200 and the car was ten years old? These are questions with which the law of contracts must deal.

Ambiguity in a collective decision is also possible with respect to contingencies surrounding an apparent collective decision. What if the temperature is so cold on the day appointed for the club picnic that picnicking would be uncomfortable? Should the members of the club understand that it is implicitly agreed that under such conditions the picnic is canceled? The possibility of rain is sometimes taken into account explicitly in such decisions, but it is never possible to account for all contingencies.

The foregoing considerations suggest the following definition of a collective decision:

> *Definition*: Collective decision: A process that identifies a pattern of future coordinated actions as the intended actions of the members of a collectivity, and creates corresponding intentions in enough members of the collectivity that in the ordinary course of events the pattern is realized.

The possible patterns of future coordinated actions that may be selected by a collective decision will generally be referred to as "options." When the options entail having different individuals fill a specific position, the options will be referred to as "candidates." The one that is selected will be referred to as the "outcome" of the collective decision or the winning candidate.

Decisions and Judgments

As Sen has pointed out, there is an important distinction between collective decisions and social welfare judgments.[3] The analysis in this book is concerned primarily with collective decisions. It deals with judgments (social welfare and other) only to the extent that they are reached through explicit collective decisions.

While a collective decision is concerned with the question, "What shall we do?" a judgment is concerned with the question, "What is best?" Very often the latter question is of interest to assist in addressing the former, but the two questions are definitely of different types. To decide what we shall do, we must coordinate our intentions. It is not necessary that we reach agreement about the relative merit of alternatives in order to coordinate our intentions. People often carry out their assigned roles in collective decisions without agreeing that the outcome is the best possible one.

There is often a coordination aspect to a judgment that is made collectively. A committee may be called upon to judge which of the essays submitted in a contest

is best. The coordination aspect arises because it will be important when the chairman announces the winner that the other members of the committee not respond with, "We decided no such thing." But no sensible person supposes that when such a decision is announced it means that the winning essay is *truly* the best, in the sense that *any* unbiased committee would necessarily have reached the same decision.

It is not possible to guarantee absolute validity in judgments no matter what process is used. Minorities are sometimes right; even a judgment that secures unanimous support may be wrong. The fact that everyone else makes a different judgment than you make may be reason for you to reexamine your judgment, but it is never a sufficient reason for changing your judgment. However, the fact that everyone else plans to *do* something different than you plan to do (such as drive on the right side of the road) is often a sufficient reason for joining the others in their plans.

Outline of the Next Five Chapters

Now that a collective decision has been defined, the next step in the analysis is to develop a taxonomy of procedures for making collective decisions. This is done in Chapter 2. Then it is necessary to identify and classify criteria for evaluating collective decision procedures. This is done in four stages: Chapter 3 reviews the criteria for evaluating changes in public policies that have been proposed by economists. With the background in ideas and terminology provided by Chapter 3, Chapter 4 offers a classification of the criteria that make collective decision procedures attractive to collectivities. Chapter 5 then evaluates the eight modes identified in the taxonomy of Chapter 2, according to the criteria developed in Chapter 4. Chapter 6 addresses the questions of how the idea of "good" applies to the eight modes of making collective decisions, and what we should understand a "good" collective decision to be.

Notes

1 Lewis Carroll, *Through the Looking Glass, and what Alice Found There* (New York, 1986), p. 70.

2 Duane M. Rumbaugh and Sue Savage-Rumbaugh, "Apes and Language Research," in Harlan B. Miller and William H. Williams (eds), *Ethics and Animals* (Clifton, 1983), pp. 213–15.

3 Amartya Sen, "Social Choice Theory: A Re-examination," *Econometrica*, 45 (1977): 53–89.

This page intentionally left blank

Chapter 2

A Taxonomy of Collective Decision Procedures

Introduction

This chapter develops the taxonomy of collective decisions procedures that is depicted in Table 2.1. The first level of classification identifies two families of collective decision procedures, those that select an option only when all members of the collectivity agree on it (the agreement-on-outcome family), and those that select an option on the basis of prior agreement to abide by a selection procedure, whether or not the collectivity achieves complete agreement on the selected option (the agreement-on-procedure family). Within each family, four modes are identified. The agreement-on-outcome family includes the modes of consensus, pseudo-consensus, trade, and extortion. The agreement-on-procedure family includes the modes of authority, random process, voting, and contest. While some collective decision procedures involve only a single mode, others involve combinations of modes.

Table 2.1 A taxonomy of collective decision procedures

The agreement-on-outcome family				
Modes:	Consensus	Pseudo-consensus	Trade	Extortion
Sub-modes:	Discovered consensus	Immediate pseudo-consensus	Purchase	Hold-up
	Achieved consensus	Achieved pseudo-consensus	Negotiated exchange	Treaty
The agreement-on-procedure family				
Modes:	Authority	Random process	Voting	Contest

The taxonomy is developed as follows. First the agreement-on-outcome family is defined formally and explained. Then the modes of this family are defined and discussed, individually and in relation to one another. The four modes of the agreement-on-outcome family are shown to lie at the corners of a square defined by the presence or absence of the properties of dissatisfaction and compensation. When one takes account of the fact that in each mode, collective decisions may be reached with or without effort, the square becomes a cube whose vertices are the eight sub-modes of the agreement-on-outcome family.

Next, the agreement-on-procedure family is defined formally and explained. The modes in this family are then defined and discussed, individually and in relation to one another. The following section offers examples of how a given decision might be made by each mode and then provides examples of collective decision procedures that combine modes. The final section discusses the relationship of reciprocity to the taxonomy.

Agreement on an Outcome

Agreement on an outcome is more primitive than agreement on a procedure, in the sense that it is necessarily used first. A collective decision can be made by a procedure from the agreement-on-procedure family only if there has previously been agreement to abide by the procedure. In the agreement-on-outcome family, on the other hand, unanimity is achieved about the coordinated actions to be taken, so that no prior agreement concerning the procedure is needed. The agreement-on-outcome family may be defined formally as follows:

> *Definition*: Agreement on an outcome: A collective decision procedure is a member of the agreement-on-outcome family if and only if the pattern of coordinated actions to be taken is determined only with the agreement of all members of the collectivity.

The next four sub-sections discuss the four modes of the agreement-on-outcome family.

Consensus

Consensus is unanimity of opinion or belief. When a collective decision is made by consensus, all members of the collectivity agree that the proposal represents their view of what the collectivity ought to do, or what they desire that the collectivity do. No one feels oppressed by the proposed outcome, nor is anyone called upon to compromise, accepting less of what he or she wants in one aspect of the outcome in exchange for others accepting less in other aspects. An example of consensus is the agreement that is achieved when a woman hopes that a man will

ask her out to dinner, and when he asks her, she accepts. Another example of consensus is provided by a jury, provided that none of its members browbeat others into agreeing on the verdict. After reaching a consensus regarding the innocence or guilt of the accused, a jury forms coordinated intentions to report their consensus when polled in the courtroom.

> *Definition*: Consensus: A collective decision is made by the mode of consensus if and only if all members of the collectivity agree that the chosen pattern of coordinated actions represents what the collectivity ought to do or what they want the collectivity to do.

The only form of persuasion that can be used to reach consensus in a collectivity is reason. It may be possible to achieve a coordination of intentions on the basis of threats or offers of compensation, but that is not consensus. Consensus may be discovered in the first exchange of views, or it may be established only after a long sequence of exchanges. The feature that distinguishes consensus from other modes in the agreement-on-outcome family is that a "meeting of minds" is achieved.

Pseudo-Consensus

Pseudo-consensus represents agreement among individuals to coordinate their actions, without agreement that the chosen pattern of coordinated actions is right or most desired. A collective decision by pseudo-consensus occurs when a man asks a woman to have dinner with him and she agrees, out of concern not to offend him, even though she does not want to go. Another example of pseudo-consensus is a jury verdict in which a jury member agrees to a verdict that he does not believe is correct, just because he does not want to argue with the other jury members. Pseudo-consensus shares with consensus the characteristic that there is no compensation for agreeing to coordinate one's actions with others.

> *Definition*: Pseudo-consensus: A collective decision is made by the mode of pseudo-consensus if and only if one or more of the members of the collectivity regard the chosen pattern of coordinated future actions as not what the collectivity ought to do and not what they desire, but all agree to coordinate their actions nevertheless, without any compensation for their agreement.

When a person agrees to an outcome that is not what he wants or what he believes the collectivity ought to do, it is usually because, in view of the orientations of the other members of the collectivity, the outcome he sees as right or most desirable appears to be infeasible. Thus the woman who reluctantly agrees to a dinner might prefer that the man involved lose interest in her, without her having to reject him, but be unable to see a way of achieving that. The juror who reluctantly agrees to a

verdict he regards as wrong might prefer that the jurors all change their minds and agree with him, but be unable to effect that.

Pseudo-consensus is distinguished from consensus by the fact that while both involve coordinating intentions, consensus entails agreement on what is right or most desirable, while pseudo-consensus does not. And unlike trade and extortion, which are discussed in the subsequent sections, no compensation is involved in pseudo-consensus.

Trade

While the search for consensus involves an attempt to achieve a coordination of intentions through a unanimity of minds, trade achieves a coordination of intentions through a process of exchange, in which individuals are promised compensation for agreeing to do things that they would not wish to do without compensation. Lawyers call such compensation "consideration," and economists often call it "bribes." A simple example of a collective decision made by the mode of trade is a purchase: If you turn over $500 to me, I will turn my car over to you. Trades are sometimes formalized in contracts. Trades can also occur in personal relationships: "I'll agree to go to the restaurant you want this week, if we can go to the restaurant I want next week."

A crucial characteristic of the mode of trade is a shared understanding that the members of the collectivity have no obligation to trade if they do not wish to do so. People are free to trade if they wish, and free to not trade if they so choose.

> *Definition*: Trade: A collective decision is made by the mode of trade if and only if some or all members of the collectivity are to be compensated for their intended actions, and if no member of the collectivity regards the way the others would have acted in the absence of a collective decision as improper.

The shared understanding that the members of the collectivity are free to refrain from trading if they so wish distinguishes trade from extortion, which is discussed next.

Extortion

While the compensation that is offered when the mode is trade involves actions that are agreed to be within the discretion of those who take them, compensation in the mode of extortion involves promises to refrain from behaving in ways that are regarded, by at least some members of the collectivity, as improper. Extortion combines the characteristic of compensation, which distinguishes trade from consensus, with dissatisfaction regarding what others are prepared to do, which distinguishes pseudo-consensus from consensus.

> *Definition*: Extortion: A collective decision is made by the mode of extortion if and only if it involves compensation and one or more members of the collectivity regard the actions that others would have taken in the absence of a collective decision as improper

Why, you might ask, is it necessary to inquire into the feelings of individuals about what others would have done in the absence of a collective decision? Suppose someone points a gun at you and says, "Give me your wallet or I will kill you." You may reach a collective decision that you will give him your wallet and he will not kill you. You are compensated with continued life for parting with your wallet. But your coordinated actions do not constitute trade, and the reason is that you regard what he would have done in the absence of a collective decision (kill you) as improper. A peace treaty is another example of a collective decision made by the mode of extortion. Agreement to release a kidnap victim in exchange for ransom is yet another. The rhetoric of labor disputes suggests that collective bargaining agreements may be as much extortion as trade. If an employer says, "If you don't want to work for the wages I'm offering, I'll find someone who will," his behavior is extortionate if you believe that it would be improper for him to hire anyone else.

The distinctions among the modes in the agreement-on-outcome family can be very subtle. A jury may think it has achieved a consensus, when some members have disagreements that they are withholding. If there are several charges against the accused, there may be explicit or implicit trade: "If you agree to acquittal on the first charge, I will agree to conviction on the second." If some of the jurors begin to browbeat others, agreement might be extorted from those who are reluctant. The distinction between pseudo-consensus and extortion can be especially difficult to make. If a person who regards the selected outcome as not proper, or not what he desires, feels that he freely chooses to coordinate his actions with the rest of the collectivity, then the mode is pseudo-consensus. But if he feels that by agreeing he secures the removal of some form of improper pressure, then the mode is extortion.

The four modes of the agreement-on-outcome family are related as follows: When a collective decision is made by the mode of consensus, the only form of persuasion that is used is reason. If the use of reason achieves agreement to coordinate future actions but not agreement on what is right or most desirable, then the mode is pseudo-consensus. If in addition to reason, offers of compensation are used to reach agreement, then the mode is trade, provided that all members of the collectivity agree that what others would have done in the absence of a collective decision would not have been improper. If agreement is reached through compensation, but some or all of the members of the collectivity feel that what others would have done in the absence of a collective decision would have been improper, then the mode is extortion.

The four modes of the agreement-on-outcome can be located at the corners of a square, in a coordinate system whose axes represent the presence or absence of the characteristics of compensation and dissatisfaction, as shown in Figure 2.1.

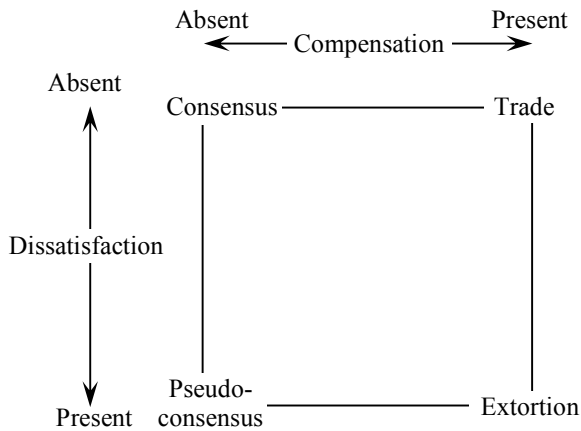


Figure 2.1 Locations of modes in the agreement-on-outcome family

For each mode in the agreement-on-outcome family, there are two sub-modes. In each case, one of the sub-modes entails agreement with minimal effort, while the other represents agreement after noticeable effort. Thus discovered consensus is consensus with minimal effort, while achieved consensus is consensus after noticeable effort. The first is represented by the example of the desired date discussed earlier, while the second is represented by a jury verdict in which true consensus was achieved after considerable deliberation. Immediate pseudo-consensus is pseudo-consensus with minimal effort, exemplified by the date to which the woman reluctantly agrees. Achieved pseudo-consensus is pseudo-consensus that emerges after considerable effort, as when a jury reaches a verdict, after much deliberation, in which some members allow themselves to be included even though they do not agree. A purchase is a trade conducted at an advertised price, where agreement is reached with practically no effort. A negotiated exchange, on the other hand, involves effort to reach agreement. A hold-up is extortion that occurs with little effort in securing agreement, while a treaty entails extortion that occurs after great effort in securing agreement.

Because the two sub-modes within each mode differ with respect to effort expended, it is possible to represent all eight sub-modes of the four modes as the corners of a cube characterized by the dimensions of dissatisfaction, compensation and effort, as shown in Figure 2.2.

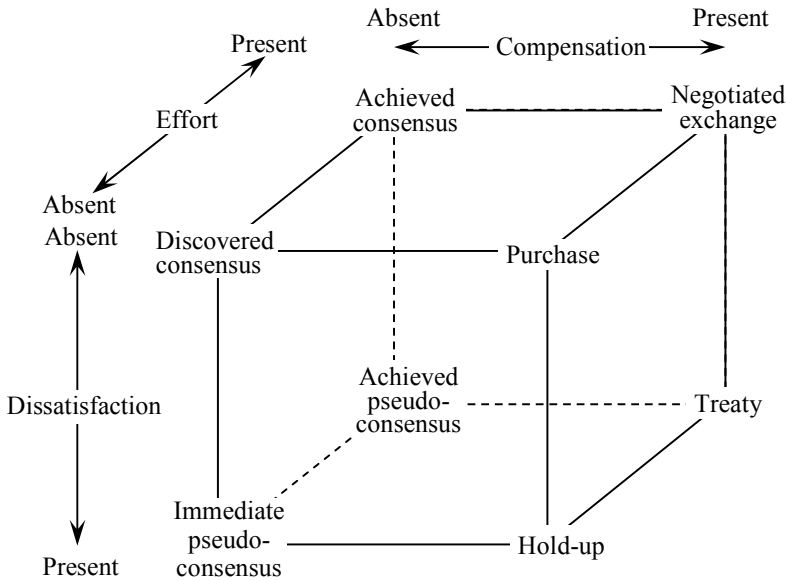


Figure 2.2 Locations of sub-modes in the agreement-on-outcome family

While the modes in the agreement-on-outcome family can be used without a prior agreement, it is also possible to formalize, through prior collective decisions, the way these modes will be used. The rituals of jury instruction, of collective bargaining, and of meetings of the Society of Friends all represent collective decisions about how collective decisions, using modes in the agreement-on-outcome family, shall be made.

Agreement on a Procedure

Agreement on a procedure represents a second form of collective decision making, in the sense that before a mode in this family can be used, there must be a prior collective decision to use it and to abide by the result that it generates.

> *Definition*: Agreement on procedure: A collective decision procedure is a member of the agreement-on-procedure family if and only if the outcome is determined on the basis of agreement to abide by a collective decision procedure.

The agreement-on-procedure family of collective decision procedures contains the four modes of authority, random process, voting and contest.

Authority

The simplest mode of collective decision-making in the agreement-on-procedure family is authority. Some person or group selects the outcome, and the whole collectivity goes along because they recognize the authority of the individual or group that makes the selection.

> *Definition*: Authority: A collective decision is made by authority if and only if, on the basis of a previous collective decision, the outcome is determined by some individual or group that does not coincide with the collectivity.

Virtually all collectivities make some of their decisions by authority. Examples would be a club president deciding the agenda for a meeting, a team captain calling a play, a wife deciding what and when the family would eat, a dictator promulgating a decree, and a seller setting a delivery date under a contract.

The authority making a collective decision need not be a single individual. A legislature, a committee, a panel of judges, and a junta are all examples of plural authorities.

When an authority is composed of more than one individual, it needs a way of making its own collective decisions. Any of the modes, including authority, can be used. A committee referring an issue to a subcommittee would be an example of authority making a collective decision by authority.

Random Process

Sometimes a random process is used to make a collective decision.

> *Definition*: Random process: A collective decision is made by random process if and only if, on the basis of a previous collective decision, a set of possible options is identified for consideration, and the outcome is selected from among these on the basis of the outcome of a random process.

One of the most common examples of a random process used to make collective decisions is flipping a coin. When there are more than two options, we may draw straws or pluck an option from a hat. A lottery was used in the U.S. not too many years ago to determine which young men would be drafted into the army. The First Book of Samuel in the Bible describes how Saul was chosen by lottery to be the first king of the Israelites (with God making sure the lottery came out right).

Random process can only be used after some other process has identified a set of options from which the outcome is to be selected. Random process shares this feature with voting.

Voting

The mode of voting is used to determine which option will prevail in an apparently rational way.

> *Definition*: Voting: A collective decision is made by voting if and only if, on the basis of a previous collective decision, a set of possible options is identified for consideration, the members of the collectivity report one or more aspects of their evaluations of these options, and an outcome is selected on the basis of a mathematical function of these reports.

The definition of voting presumes that the set of voters coincides with the collectivity making a decision. If the set of voters is either larger or smaller than the collectivity, then an element of authority enters into the decision. For example, suppose that the voters of a state pass a referendum that affects what will be taught in public schools. The voters are exercising authority over teachers even if every teacher is a voter. And when the legislature passes a law that affects all voters, the legislature is exercising authority. Thus the body that exercises authority can be either larger or smaller than the set of persons whose intentions are coordinated by a collective decision.

The evaluations of options under voting may take a variety of forms. The most common form is for each voter to report which of the specified options he or she considers best. Another common form is for each voter to submit a ranking of the specified set of options. Yet another is for each voter to give a score to each of the specified set of options, according to a prescribed scale. Another form that has been suggested is that each voter should report whether he or she "approves" of each option. Yet other possibilities will be introduced later.

When there are more than two options, there is no way of insuring that voters are motivated to report accurate evaluations. This is proved in Chapter 11, and ways of nearly overcoming this difficulty are presented in Chapter 16.

There are many ways in which the evaluations of voters can be combined. Chapters 13 through 16 explore these in great detail.

Contest

Ancient records of the laws of virtually all the tribes that settled in Europe include provision for a wide variety of civil and criminal disputes to be settled by armed contests between accuser and accused or their champions.[1] As late as 1818, such a provision was employed in an English murder case. After a jury acquitted the

defendant, Thornton, the victim's brother, Ashford, appealed. Thornton offered to defend himself against the charge by engaging in battle with Ashford, and when Ashford withdrew from the contest Thornton was released.[2]

A general rationale for making collective decisions by contest is that the relative value of options can be discerned by pitting them against one another in a procedure that requires effort or sacrifice by their proponents.

> *Definition*: Contest: A collective decision is made by the mode of contest if and only if, on the basis of a previous collective decision, an outcome is selected on the basis of the ability of the advocates of the outcome to surpass the advocates of other options in some endeavor involving effort or sacrifice.

Contest is a somewhat elusive mode, because it occurs most commonly in conjunction with other modes. Before an election that is decided by voting, a contest is often held for the votes of the voters. Before a judge makes a decision by authority or a jury makes a decision by consensus, a contest for his mind or theirs is held in the courtroom. Before a contract is let, a contest may be held for the lowest or best bid. In these instances the collective decision that is made may be described as one of who won the contest. Contest occurs in its pure form when there is no need for an arbitrator to decide who won the contest, because there is no possible ambiguity.

Examples of the Modes and their Mixtures

As an exercise in the understanding of the different modes, consider how the question of what restaurant a group of people will eat at might be decided in each of the different modes:

Consensus: "How about the Lotus Garden." "Sure." "Fine. "Agreed."

Pseudo-consensus: "I'd like to go to the Lotus Garden." "If you insist."

Trade: "I don't particularly like Chinese, but I'll go along if we can go to a Mexican restaurant next time." "OK, it's Mexican next time."

Extortion: "I can't stand Chinese. If you want me to go along, you'll have to go somewhere else. In fact, Sam's is the only decent restaurant in this town." "OK, we don't want a row. We'll go to Sam's."

Authority: "Bill, it's your turn to choose the restaurant. Where shall we go?"

Random process: "It's between French and Italian. Shall we flip a coin?"

Voting: "We have three votes for Polynesian and two for Greek, so Polynesian it is."

Contest: "The Golden Dragon came in with the lowest bid, so we'll be having our banquet there."

While each mode of making collective decisions can be defined in such a way as to distinguish it from the others, actual collective decisions often entail either

mixtures of modes or ambiguity as to what mode was used. The following paragraphs provide examples, for each of the 28 possible pairs of modes, of collective decisions that are either mixtures of the two modes or ambiguous cases that could be either mode.

Consensus and pseudo-consensus: When a jury has deliberated a long time, it is often difficult to know whether the outcome is consensus or pseudo-consensus.

Consensus and trade: When consensus is sought and agreement is reached through compromise (for example, a couple deciding where to take their vacation), it is sometimes difficult to say whether the result represents a consensus or a trade.

Consensus and extortion: As the jury room door closes, one member of the jury says, "If you don't want this jury locked up for days on end, you'll all vote for acquittal." When the jury is polled by the foreman, they all announce that they favor acquittal.

Consensus and authority: When consensus is sought and a person who has authority over the collectivity in other matters firmly asserts that there is only one suitable outcome, any resulting agreement could be an expression of either consensus or authority. (Concern for this led to the custom in the U.S. Supreme Court of having the Associate Justices express their views in the Court's discussions in order of reverse seniority, with the Chief Justice going last.)

Consensus and random process: In the middle of a contentious effort to reach consensus, a peal of thunder is heard as one person speaks. This is taken by all as an indication that the outcome favored by the speaker is the right one.

Consensus and voting: When a jury is supposed to deliberate until they reach a unanimous verdict, there is a temptation, after a poll is taken, for those in the minority to think that they ought to concede simply because they are in the minority. If the fact that they are a minority leads them to believe that the majority is right, then it is not clear whether the decision was made by consensus or voting.

Consensus and contest: A panel of judges is told to reach a consensus about who should receive the first prize for apple pie at a county fair.

Pseudo-consensus and trade: A man tells a woman he is taking her to dinner. She knows from past experience that he will regard her to be obliged to spend the night with him in exchange for the meal. He does not realize that she places no value on the meal, and that she acquiesces reluctantly to his domineering ways because, despite his boorishness, she likes him (he resembles her father), and she does not want to hurt his feelings.

Pseudo-consensus and extortion: A teen-ager asks for use of the family car. After initially resisting, the father agrees when he realizes that if he does not agree, his moping child will make it completely impossible for him to enjoy the evening.

Pseudo-consensus and authority: An adult child continues to behave as a parent directs, while resenting the direction.

Pseudo-consensus and random process: In the middle of an argument between two persons, one person says, "The only way to settle this is to flip a coin." He

then flips a coin and asks the other person to call it. The other complies resentfully, while feeling that the matter should have been settled by reason.

Pseudo-consensus and voting: In a club, a motion is made and seconded. In the ensuing discussion, several members express strong reservations about the motion, but when a vote is taken, no one votes against it.

Pseudo-consensus and contest: Two children fight over whose turn it is to ride a bicycle. One says, "I'll let you use it if you can beat me in a race around the block." The other accepts the challenge, despite believing that his claim to the bicycle was better.

Trade and extortion: When a person grumbles about high prices and then buys something, it is not clear whether the mode is trade or extortion.

Trade and authority: If a baseball player and the team for which he plays are unable to agree on what his salary ought to be, the dispute is referred to an arbitrator. Each party proposes a salary, and the arbitrator chooses between the two proposals.

Trade and random process: Trades in over-the-counter stocks are arranged by brokers in telephone conversations. Inevitably, misunderstandings occur from time to time. If the parties to a misunderstanding are unable to agree on what happened, one of them flips a coin, the other calls it (by phone), and the one who flipped the coin announces the result.

Trade and voting: Logrolling is the classic example of this combination.

Trade and contest: A contract may go to the bidder who offers the best combination (according to a specified schedule) of low price and prompt delivery.

Extortion and authority: When a person in authority is blackmailed into making a particular decision, the result is a combination of extortion and authority.

Extortion and random process: Casino owners prefer not to permit bets by blackjack experts who keep track of which cards are used and thereby identify favorable bets. Considering the practice not legitimate but being obliged by law to serve all customers, owners have been known to pay off consistent winners not to play at their establishments. A person with very good luck might feign such expertise even though he did not have it, and thereby be paid off not to play.

Extortion and voting: Intimidation of voters exemplifies this combination.

Extortion and contest: Halfway through a race to determine which of two suitors will have the hand of a lady, the leader loses a shoe. He tackles his rival, and refuses to let him go until he has removed a shoe too.

Authority and random process: When a seer is called in to read entrails he is supposed to let the random selection of the beast determine the outcome. But if he thinks he knows what the answer ought to be, he can cheat, and since no one else can read entrails, he will get away with it. Since everyone knows this, they may expect and even count on the seer to do so. (Is there a parallel with macroeconomic forecasting?)

Authority and voting: In the State of Virginia, when there is a tie vote in a county Board of Supervisors, a local judge appoints someone to break the tie.

Authority and contest: A single judge of an essay contest exemplifies this combination.

Random process and voting: When an election for a public office results in a tie, the most common way of determining the winner is to use some random process.

Random process and contest: This combination would be exemplified by choosing a club president according to the result of a competition involving both chance and skill, such as playing solitaire.

Voting and contest: A debate before an audience, followed by a vote, exemplifies this pair.

The Relationship of Reciprocity to the Taxonomy

While the taxonomy has been used to classify many collective decisions, one phenomenon that has not come up in the discussion so far is reciprocity. One might wonder whether the taxonomy can classify this phenomenon. Reciprocity is the granting of favors in the expectation that favors will be returned, should the need arise. Neighbors providing "a cup of sugar" for each other is a classic example of reciprocity. Politicians supporting each others' interests is another. Reciprocity has something in common with trade, except that the compensation is unspecified, and may not occur at all. Each occurrence of reciprocity is a consensus or pseudo-consensus, since compensation is not provided. A decision by a group to behave with reciprocity toward one another could be either trade or consensus, depending on whether the members of the group see their deference as something that requires compensation, or something that is proper without regard to compensation. Reciprocity does not have a unique location in the taxonomy, because it entails a sequence of collective decisions, by different modes, rather than a single collective decision. But any individual collective decision in a framework of reciprocity can be classified using the taxonomy.

Now that a taxonomy of collective decision procedures has been provided and its usefulness has been shown, the next step in the analysis is to develop criteria for evaluating collective decisions. The next chapter discusses criteria that economists have proposed for evaluating public actions. Discussion of these provides a background for a more general inquiry in the subsequent chapter into criteria by which collective decision procedures might be evaluated.

Notes

1 Henry Charles Lea, *The Duel and the Oath* (Philadelphia, 1974), esp. pp. 107–108.

2 Lea, *The Duel and the Oath*, p. 246.

This page intentionally left blank

Chapter 3

Economic Criteria for Evaluating Collective Decisions

Introduction

This chapter discusses four ethical traditions of twentieth century economics, associated with the names Pareto, Bergson–Samuelson, Hicks–Kaldor, and Buchanan–Tullock.

Economics originated in the eighteenth century as political economy, which inquired into good political structures for dealing with the economic organization of a nation. Economics and political economy have thus been concerned with evaluating the collective decisions of political entities.

The principal ethical framework of political economy in the nineteenth century was utilitarianism: What is good is what provides "the greatest good for the greatest number." There was a tendency to treat this idea as if its meaning were self-evident and did not require a rigorous specification.

An important component of nineteenth-century political economy was the theory of how prices are determined. Progressively more sophisticated theories of price determination led to growing concern for the "paradox of value," the puzzle that things like water that were obviously of great value because they were essential for life seemed to command unreasonably low prices, while things like diamonds that were clearly not essential could command remarkably high prices. In the 1870s, three economists, Jevons, Menger, and Walras, arrived independently at the understanding that the price of an item was related to its "marginal utility" (the utility to a purchaser of the last unit of the item purchased) and not to its "average utility" (the impact on the utility of a purchaser of being able to purchase the good, divided by the number of units purchased). This understanding has since become known as the "marginal revolution."[1] With the marginal revolution came a much greater focus on the explanation of individual behavior and a change in the name of the discipline from "political economy" to "economics."

In the theories that developed from the marginal revolution, economic behavior is treated as determined by a relatively small number of parameters: A person's purchases of bread are determined by his "demand schedule for bread." If you know a person's demand schedule for bread and you also know the price of bread, the prices of goods that are substitutes for bread, the prices of goods that are

complements of bread, the person's income, and the price level, then you can determine what the person's purchases of bread will be. For analytic purposes, a person *is* his demand schedules, or, as the theory developed in the early Twentieth Century, his "utility function." All of a person's purchases are determined, according to modern economic theory, by a person's utility function, his income, and the prices of all goods.

Pareto

The idea of representing people as utility functions led Vilfredo Pareto to an innovation in the ethics of economics, near the turn of the Twentieth Century. In Pareto's words (translated from French),

> Consider any position and assume that we move away from it by a very small amount, consistent with the restrictions [on what is economically feasible]. If in so doing the welfare of all the individuals of the collectivity is increased, it is obvious that the new position is more advantageous to each of them; and vice versa, it is less so if the welfare of all of the individuals is decreased. Moreover, the welfare of some of them can remain the same, without changing these conclusions. But on the other hand, if this small movement increases the welfare of certain individuals and decreases that of others, we can no longer state positively that it is advantageous to the entire collectivity to carry out that movement.
>
> These considerations lead to defining the position of maximum ophelimity [a word introduced by Pareto, meaning "economic utility"] to be one from which it is impossible to move a very small distance, in such a way that the ophelimities of the individuals, except for some which remain constant, all increase or decrease.[2]

The first paragraph defines what came to be known in economics as a "Pareto improvement," a change that makes one or more persons better off and no one worse off. The second paragraph (less the last two words, which make no sense) defines what came to be known as a "position of Pareto optimality," a position from which it is not possible to make a Pareto improvement.

The meaning of "position" in these definitions is an allocation of goods to all members of a society. The meaning of "better off" is "at a higher level of utility, as determined by the person's utility function." In the twentieth century, the concepts of Pareto improvement and Pareto optimality have been more widely employed by economists than any other ethical concept. The application of Paretan ethics to collective decisions is supported by the slogan, "You can't make interpersonal comparisons of utility."

Discussion

Recall that the idea of a "Pareto improvement" came from Pareto's concept of a movement from some initial position, such that the welfare of no member of a collectivity is decreased, while the welfare of some is increased. The idea of an increase in a person's welfare is much more complex than might be apparent.

If, by a movement that increases the welfare of an individual, Pareto had meant a decision to which the individual consents, then a Pareto improvement would have been equivalent to a collective decision made by consensus. But Pareto's use of the concept of "improvement in welfare" is different.

Pareto treated the preferences of individuals as known. With such an assumption, it is not necessary to consult individuals to see whether they consent to any proposed decision. Whether it will increase or decrease an individual's utility can be determined by consulting the individual's utility function. Thus, from one perspective, the idea that a change is good only if it is a Pareto improvement is a recognition of the freedom of individuals to choose what they wish. But from another perspective it imposes, on all individuals, decisions made by an omnipotent, omniscient and benevolent authority. It should be noted, however, that the Paretan framework greatly constrains the authority making collective decisions. Only decisions that harm no one are endorsed.

There is another conceptual difficulty with the Paretan framework that should be mentioned. In focusing on a "movement" from some initial "position," Pareto implicitly assumes that it is possible to specify what the initial and subsequent positions are, and that it is possible to implement a decision to move from the former to the latter. At the time when Pareto wrote, this might not have seemed as strong a set of assumptions as it seems to be from a later perspective, because in Pareto's era there was more of a tendency than there is today to omit from formal economic theory the consideration of future consequences of possibilities that were analyzed. Thus a "position" would be a description of the state of an economy at a single moment in time, and the movement analyzed by Pareto would be to an alternative description at the same moment in time.

While some economic analysis is still conducted in this "comparative statics" framework, there is a growing tendency in economics to recognize the need to incorporate into economic analysis consequences that emerge over time. From such a perspective, a Pareto improvement is a movement of an economy from one path through time to another. In principle, the path would extend to the furthest reaches of time, and at each point in time all productive activities and all goods consumed by each economic actor would be specified. It is clearly not possible to specify either a "position" or a "movement" in these terms. Apart from the problem of specifying all economic events, there are unavoidable changes that arise as people are born and die, and as resources are exhausted. Even if by some Herculean effort the initial position and the movement could be specified, the concept of a Pareto improvement faces the further difficulty that it is inconceivable

that any human institution could insure that the movement would occur as specified. Thus "Pareto improvement" is a theoretical construct without an operational counterpart.

From a perspective somewhat different from Pareto's, there is nevertheless practical value in the idea of Pareto improvements. People often envisage the future as a replication of existing conditions into the indefinite future, even if this is not feasible. Suppose people agree that this defines the initial "position," and they accept Pareto's dictum that a change is an improvement only if no one is made worse off. If they make their decisions by consensus, they may be able to use these ideas to reach consensus more quickly. If decisions are made by authority, the authority will be assisted and also constrained in ways that are likely to be attractive to collectivities. Even if true Pareto improvements are not found, "approximate Pareto improvements," that give approximately suitable compensation to all broad classes of persons affected by a change may be identified.

A limitation of this approach arises when the unfolding of time reveals that an initial pattern cannot be replicated indefinitely: Oil from the ground becomes available only at a greater resource cost, and someone must be disappointed. The Paretan framework gives every person a reason for saying that he should not have to suffer disappointment. Suppliers believe that prices should rise when costs rise, while buyers see any rise in price as an infringement on their rights. It is not possible to adhere to the general rule that no change is made that makes some persons worse off. Conflict is often the result.

An additional limitation of the Pareto principle arises with respect to the rectification of past wrongs. The principle incorporates implicitly the idea that if anyone has illicitly initiated a change that is not a Pareto improvement, then this change should be undone. But what is the starting date for rectification? If yesterday's thieves must return their ill-gotten gains, what about the thieves of last year or the last century? The Pareto principle can be implemented only after a collectivity has agreed on a time from which changes are measured. The difficulty of reaching such an agreement is illustrated by current conflicts in the Middle East and in Northern Ireland. There are Palestinians who carry around their necks the keys to houses their families left in 1948. There are Catholics in Northern Ireland who can identify the land from which their families were driven in the 17th Century. A commitment to the Pareto principle would require either a way of rectifying all such wrongs, or a consensus about which ones need not be rectified.

Bergson–Samuelson

Pareto's formulation fits with the impetus in economics toward greater rigor, but it was dissatisfying in its extreme agnosticism with respect to the evaluation of proposals that make some persons better off and others worse off. A rigorously

defined concept of "good" that was not so agnostic was provided by the combined efforts of Abram Bergson and Paul Samuelson. In a 1938 article, Bergson summarized the mathematical conditions required for the maximization of "economic welfare" conceived as a mathematical function.[3] Bergson gave special emphasis to the case in which the marginal utility of additional income is the same for all persons when they have the same income, while also noting that, "in general, any set of value propositions that is sufficient for the evaluation of all alternatives may be introduced, and for each of these sets of propositions there corresponds a maximum position."[4]

Paul Samuelson elaborated and clarified Bergson's analysis.[5] With respect to the value judgments required for a social welfare function, Samuelson wrote,

> Without inquiring into its origins, we take as a starting point for our discussion a function of all the economic magnitudes of a system which is supposed to characterize some ethical belief—that of a benevolent despot, or a complete egoist, or "all men of good will," a misanthrope, the state, race, or group mind, God, etc. Any possible opinion is admissible, including my own, although it is best in the first instance, in view of human frailty where one's own beliefs are involved, to omit the latter. We only require that the belief be such as to admit of an unequivocal answer as to whether one configuration of the economic system is "better" or "worse" than any other or "indifferent," and that these relationships are transitive; i.e., *A* better than *B*, *B* better than *C*, implies *A* better than *C*, etc.[6]

Samuelson then notes that the question of which aspects of the economy could conceivably be varied and which are immutable is a question of the level of abstraction at which the analyst chooses to work.[7] He then introduces a variety of constraints that an analyst might wish to place on the ranking of alternative configurations of the economic system.[8] One of the more critical of these possible constraints is the fifth,

> … that individuals' preferences are to "count." If any movement leaves an individual on the same indifference curve, then the social welfare function is unchanged, and similarly for an increase or decrease.[9]

This assumption permits the social welfare function to be expressed as

$$W = F(U_1, U_2, \ldots),$$

where the U_i's are individual utility functions, each of which is a function of the goods consumed by that individual. It is this formulation that has come to be known as the Bergson–Samuelson social welfare function.

It is implicit in this framework that collective decisions will be made by the mode of authority, with government officials as the authorities. The role of the social welfare function may be conceived as one of facilitating communication

between government officials and economists who advise them. The Bergson–Samuelson framework suggests that when a policy is under consideration that will be advantageous to some members of a society and disadvantageous to others, an economist who is advising government officials should ask the officials to specify the social welfare function (the way in which the welfare of the society varies with the utility levels of the members of the society). With this information, the economist would then employ economic tools to determine the impact of the proposed policy on social welfare.

Social welfare functions are widely used in economic theory, but they are seldom if ever used as the basis for actual collective decisions. This is probably because, true to Pareto's dictum, no one can find an acceptable basis for making the judgments about the relative satisfaction that different individuals receive from additional income.

Discussion

The Bergson–Samuelson framework avoids many of the difficulties of the Paretan framework, by not proposing to make no one worse off. It cuts through the problem of comparing the satisfactions of different individuals by assuming that the analyst has at hand a function that provides a transitive ranking of all economic possibilities that might be encountered, and that the analyst is prepared to credit that function as valid. This makes it possible for the analyst to compare any two allocations of goods to individuals and determine which of the two is better in terms of the specified social welfare function. However, the problem of how one would actually obtain a ranking of economic possibilities that deserved to be credited with ethical validity remains. The implicit suggestion of the Bergson–Samuelson framework is that the government has all the answers. But if one is a government official who is supposed to have the answers and wonders from where to get them, or if one has doubts about the answers offered by government officials and wishes to appraise them by some standard, the Bergson–Samuelson framework offers no guidance.

Hicks–Kaldor

The principal framework actually used by economists who evaluate government policies is "cost–benefit analysis." From one perspective, cost–benefit analysis represents a highly simplified version of a Bergson–Samuelson social welfare function: Suppose that all members of a society always derive the same satisfaction from additional income, and anyone's loss of satisfaction from having one dollar less is equal to his gain in satisfaction from having one dollar more. Then the welfare of the society (in Bergson–Samuelson terms) is maximized when the aggregate income of the society is maximized. Thus under the assumption of

equal satisfaction from additional income, any policy can be evaluated by comparing the gains of those who gain with the losses of those who lose, all measured in money.

A different intellectual justification for cost–benefit analysis originates in a pair of articles by John Hicks and Nicolas Kaldor in 1939.[10] Their idea was that if it was necessary to redistribute income to maximize the economic welfare of a society, then this could always be done separately from allocative decisions, which could therefore be made solely on the basis of maximizing aggregate value measured in money. As Kaldor put it,

> In all cases, therefore, where a certain policy leads to an increase in physical productivity, and thus of aggregate real income, the economist's case is quite unaffected by the question of the comparability of individual satisfactions; since in all such cases it is *possible* to make everybody better off than before, or at any rate to make some people better off without making anybody worse off. There is no need for the economist to prove—as indeed he never could prove—that as a result of the adoption of a certain measure nobody in the community is going to suffer. In order to establish his case, it is quite sufficient for him to show that even if all those who suffer as a result are fully compensated for their loss, the rest of the community will still be better off than before.[11]

Hicks and Kaldor said, in other words, that a policy is good if it provides a "potential Pareto improvement," a change that would be a Pareto improvement if appropriate compensation were paid. It need not actually be paid, they argued, because if it were necessary that the compensation be paid for the policy to increase social welfare, then a decision could be made to pay it.

Discussion

The Hicks–Kaldor framework sidesteps the problem of the Bergson–Samuelson framework, by asserting that distributive issues can be handled separately from allocative ones. In terms of the Bergson–Samuelson framework, this suggests that one must be making one of the following three assumptions:

1. The authority responsible for making collective decisions has the capacity to implement costlessly whatever measures are necessary to undo the adverse distributive consequences of any allocative decision.
2. The social value of additional income for different persons is always the same.
3. There are no distributive consequences of allocative decisions.

The first of these assumptions is the one that Hicks and Kaldor appear to have made, implicitly. To state it explicitly is to reveal its implausibility. While economic theory can manipulate systems in which the preferences of all

individuals are known, the problem here is to *ascertain* the consequences of alternative actions upon individuals, and then *act* on the knowledge acquired. There are procedures that motivate individuals to reveal honestly the impacts of alternative actions upon them, under particular rules about how the information so acquired affects collective decisions and what payments are required of individuals.[12] However, these procedures do not provide a motivation for honest statements of consequences if the information will be used to undo the distributive effects of the choices. We do not have a way of doing what Hicks and Kaldor assume we can do.

The second assumption is quite indefensible as an ethical assertion. If anyone thought it might be valid, critics of Hicks and Kaldor pointed out that it entails problems of consistency that arise as a consequence of the difference between payments that people are willing to make to achieve changes they desire and the compensation they require to accept changes that they would rather not have.[13] Consider the problem of allocating a week's vacation in a house and in a cottage to two people who would both rather have the house. It would not be at all implausible for both people to require compensation of $100 to accept the cottage instead of the house, while being willing to pay only $80 to have the house instead of the cottage. That is, a person who has the house is indifferent between keeping the house and accepting $100 to move to the cottage, while a person who has the cottage is indifferent between keeping the cottage and paying $80 to move to the house. Economists describe the difference in the amounts as caused by an "income effect." When a person has more (the house instead of the cottage) he is generally willing to pay more for a given change in his circumstances than when he has less. But this means that whichever way the house and the cottage are allocated, changing to the other allocation produces a net loss of $20, if one simply aggregates the willingnesses to pay. Since the cost of making one change in allocation and then the other must be 0 (assuming that they remain at the hypothetical level, in the mind of the analyst, and are not actually implemented), it must be the case that the social value of additional income to all persons is not always the same.

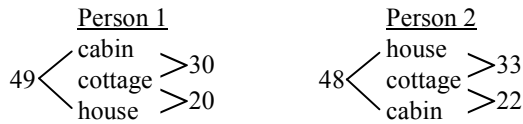
One feature of this paradox that makes it less than completely devastating is that it promotes not acting, leaving things as they are. The consequences of a paradox that asserted that both moving from A to B and moving from B to A produced gains would be more serious. While this does not occur with just two options, there are examples for three options, where the summation of dollar consequences implies that A is better than B, B is better than C, and C is better than A.

Suppose that two persons must decide whether to spend a week's vacation together in a cabin at a cost of $50 each, a cottage at a cost of $100 each, or a house at a cost of $200 each. The first person, being frugal, ranks the options cabin, cottage, house, and would pay $30 to have the cabin rather than the cottage with its expense, $20 to have the cottage rather than the house with its greater

expense, and \$49 to have the cabin rather than the house (not \$50, because of an income effect). The second person, being a Sybarite, ranks the options house, cottage, cabin, and would pay \$33 to have the house rather than the cottage, even with greater cost of the house, \$22 to have the cottage rather than the cabin, and \$48 to have the house rather than the cabin (not \$55, because of a stronger income effect). In the top panel of 3.1 below, the valuations are shown as numbers connecting pairs of options. The bottom panel of Example 3.1 shows the sums of the valuations. These indicate that the cottage is more valuable than the cabin, the house is more valuable than the cottage, and the cabin is more valuable than the house. Thus the idea that the sum of individual valuations provides a reliable indication of an increase in value is disproved.

Example 3.1 Two persons evaluating three options

Individual valuations of moves



Combined valuations of moves

Person	Cabin to cottage	Cottage to house	House to cabin
1	–30	–20	49
2	33	22	–48
Total	3	2	1

The third possible assumption, that allocative decisions do not have distributive consequences, seems the most implausible of all. And yet, there is a perspective from which it makes some sense. Polinsky has suggested that if the cost–benefit rule is evaluated over many applications, if there is a net benefit with each application, and there are no systematic correlations between who happens to lose in one application and who happens to lose in another, then the law of large numbers should come close to guaranteeing that all persons benefit approximately equally from all applications taken together.[14] This is the most reasonable defense of the Hicks–Kaldor approach. For this defense to work, it is essential that there be many applications of the cost–benefit criterion and that there be no systematic correlations between the losers in one application and the losers in another.

Buchanan–Tullock

A fourth ethical approach is used in the book, *The Calculus of Consent*, by James Buchanan and Gordon Tullock. They say,

> When will it prove desirable to shift one or more sectors of human activity from the realm of private to that of social choice, or vice versa? Implicit in our discussion is the assumption that the criteria for answering such questions as this can only be found in the conceptual unanimity among all parties in the political group. Agreement among all individuals in the group upon the change becomes the only real measure of "improvement" that may be accomplished through change.[15]

This idea has a close connection to the ideas of Pareto. However, where Pareto focuses on evaluating alternative allocations, Buchanan and Tullock focus on evaluating alternative rules or institutions. Also, where the Paretan criterion focuses on what the analyst can deduce from detailed knowledge that he assumed to have about the preferences of individuals, the Buchanan–Tullock criterion suggests (though there is some ambiguity on this point) that their concern is with actual agreement.

Discussion

Buchanan and Tullock identified improvement with conceptual unanimity. The philosopher John Rawls made a similar suggestion a few years earlier:

> The idea is that everyone should be required to make *in advance* a firm commitment, which others also may reasonably be expected to make, and that no one be given the opportunity to tailor the canons of a legitimate complaint to fit his own special condition, and then to discard them when they no longer suit his purpose. Hence each person will propose principles of a general kind which will, to a large degree, gain their sense from the various applications to be made of them, the particular circumstances of which being yet unknown.[16]

In the Rawls–Buchanan–Tullock framework, the cost–benefit criterion can be defended on the ground that it is a rule for governmental decisions to which people can be expected to agree before they know what their particular interests will be in the cases to which the rule will be applied. By focusing on rules and separating agreement on rules from agreement when the rules are applied, Rawls, Buchanan and Tullock avoid the need to justify the redistributive consequences of any particular application of the accepted rules. The insight of Rawls, Buchanan and Tullock parallels the idea of using a collective decision procedure from the agreement-on-procedure family, so that it will not be necessary to achieve agreement on all outcomes.

A troublesome aspect of the Rawls–Buchanan–Tullock framework is its ambiguity with respect to whether their criterion is one of hypothetical, "conceptual unanimity" or one of actual agreement. Here it is necessary to distinguish between, on the one hand, their presentation of their conception of an ideal government and why they favor it, and, on the other hand, the possible use of their framework to justify disputed actions of actual governments.

For the purpose of presenting and defending a conception of an ideal government, an argument of the form, "Rule X is attractive because it is something that people could be expected to agree upon," should be an acceptable form of argument. Conceptual agreement is all one can require of a proposal, because at the time of the proposal there is no possibility of actual agreement. On the other hand, if there is actual disagreement about what a collectivity ought to do, it would be highly suspicious to say, "People would agree on rule X if they were in circumstances conducive to reaching agreement. Therefore it is acceptable to impose rule X on people who say they do not want it and have never said they want it." For the defense of actual governments, the spirit of the Rawls–Buchanan–Tullock framework requires not conceptual agreement but actual agreement. The practice of searching for rules that satisfy a criterion of conceptual agreement is valuable, nevertheless, because it may lead to the identification of rules that are able to secure actual agreement.

Summary

Pareto identified a good change as one that made at least one person better off and no one worse off. This has much in common with the idea of making collective decisions by consensus or trade. But Pareto ignored the problem of how one ascertains whether people are actually better off, as well as the problems of how one identifies "the initial situation" in a dynamic world, and how one ensures that a specified change will occur.

The framework advanced in this book, by contrast, does not ask whether people are better off. It focuses instead on the agreements to coordinate intentions, recognizing the possibility that these agreements may be reached through the modes of pseudo-consensus or extortion. In this framework there is no need to identify "the initial situation." The focus on intentions also highlights the possibility of discrepancies between what is planned and what actually occurs.

Bergson and Samuelson identified a good change as one that is approved by the authority who specifies the social welfare function. This framework may be useful as a way for economists to interact with political leaders who hire them, but it provides almost no guidance to anyone who wishes to evaluate the existing political process rather than concede all authority to it.

Hicks and Kaldor identified a good change as one such that the gains of those who gained are greater than the losses of those who lost, when all gains and losses

are measured in money. This might be characterized as dollar-weighted voting (although the "voting" would be conceptual rather than actual, as the evaluation would be undertaken by an analyst rather than by a collective decision procedure). The virtue of the Hicks–Kaldor rule rests either on an assumption that redistribution can be implemented costlessly or on an assumption that all persons place the same value on additional money, and neither of these assumptions is tenable.

Buchanan and Tullock (and Rawls before them) identified a good rule as one that people would agree upon if they could not predict what their own circumstances would be when the rule was applied to them. This is akin to the idea of using collective decision procedures from the agreement-on-procedure family, so that agreement on outcomes is not necessary.

This completes the survey of how economists have evaluated collective decisions. The next chapter discusses general criteria for evaluating collective decision procedures.

Notes

1 Mark Blaug, *Economic Theory in Retrospect* (Homewood, 1962), pp. 272–3.
2 Vilfredo Pareto, *Manual of Political Economy* (New York, 1971), pp. 451–2.
3 A. Bergson, "A Reformulation of Certain Aspects of Welfare Economics," *Quarterly Journal of Economics*, 52 (1938): 310–34.
4 Ibid., p. 323.
5 Paul Samuelson, *Foundations of Economic Analysis* (Cambridge, 1947), pp. 203–53.
6 Ibid., p. 221.
7 Ibid.
8 Ibid., pp. 222–8.
9 Ibid. p. 223.
10 N. Kaldor, "Welfare Propositions in Economics," *Economic Journal*, 49 (1939): 549–52: J. R. Hicks, "Foundations of Welfare Economics," *Economic Journal* 49 (1939): 696–712.
11 Kaldor, "Welfare Propositions in Economics," p. 550.
12 See Chapter 16.
13 T. Scitovsky, "A Note on Welfare Propositions in Economics," *Review of Economic Studies* 9 (1941–42): 77–88.
14 A. Mitchell Polinsky, "Probabilistic Compensation Criteria," *Quarterly Journal of Economics* 86 (1972): 407–25.
15 James Buchanan and Gordon Tullock, *The Calculus of Consent* (Ann Arbor, 1962), p. 6.
16 John Rawls, "Justice as Fairness," *Philosophical Review*, 57 (1958): 171–2.

Chapter 4

General Criteria for Evaluating Collective Decision Procedures

Introduction

Why do people make collective decisions by one procedure rather than another? A collectivity's first decision is necessarily made by a mode from the agreement-on-outcome family. In most cases, this first decision involves voluntary participation, and the procedures for making further collective decisions are then chosen deliberately and collectively, from a range of possibilities. On the other hand, it is possible for a collectivity's first decision to involve coercive participation, subjugating one individual or group to another, in which case subsequent collective decisions can represent the exercise of authority by the subjugator over the subjugated. Then the choice of a mode expresses the exercise of power rather than pursuit of the common good. But such cases are exceptional. Thus in most cases collective decision procedures are chosen because they satisfy criteria that make them more attractive than their alternatives. This chapter explores a taxonomy of these sometimes conflicting criteria. The taxonomy is shown in Table 4.1.

Table 4.1 A taxonomy of criteria for evaluating collective decision procedures

Efficiency		Equity			Stability	
outcome efficiency	procedural efficiency	symbolic equality	material equality	productivity-based inequality	restoration	preservation

It might be argued that the selection of a procedure determines the outcomes that will be selected through its use, and therefore collectivities do not apply criteria to procedures, but rather to outcomes implied by procedures. While there is an element of validity to this argument, it is usually not possible to foresee precisely what outcomes are implied by the choice of a procedure. Thus procedures are evaluated on the basis of generalizations about their likely consequences, with the outcomes that will be chosen remaining conjectural. The explanation of preferences for collective decision procedures therefore involves entering the imagination of a person confronted with the uncertain prospects of different

procedures and asking what criteria he or she would apply. This method has much in common with that of Buchanan and Tullock in *The Calculus of Consent*.

Efficiency

The appeal of efficiency is the appeal of achieving more of whatever people want. More can be achieved either by increasing outcome efficiency, that is, by increasing the overall attractiveness of selected outcomes, or by increasing procedural efficiency, that is, by using a less costly procedure for selecting outcomes.

Ideally, the efficiency of an outcome would be measured by summing the amounts of utility that individuals gain and lose by having that outcome instead of a specified alternative. However, since there is no way of measuring individual utility that yields incontrovertible results, economists generally measure the efficiency of an outcome by summing the amounts of money that the members of a collectivity would be willing to give up, or would need to be paid, to have that outcome instead of a specified alternative.

For example, if in a household of seven persons there are three who would value a one-hour postponement of the time when dinner is served at $5 each, while the other four would experience a cost of $3 each for the postponement, then an economist would say that it is efficient to postpone the time when dinner is served, because the benefits ($15) exceed the costs ($12). This calculation makes no allowance for the possibility that money is more valuable to some persons than to others. This accommodation, in the measurement of efficiency, to what is feasible means that measured outcome efficiency is a less pertinent concept than might be desired, but it still has a significant bearing on the attractiveness of collective decision procedures.

Procedural efficiency is measured, inversely, by the magnitude of the resource and psycho-social costs of making decisions. One important resource cost of making collective decisions is the time of members of the collectivity. Time spent making decisions generally has valuable alternative uses. There may also be transportation costs in bringing the members of the collectivity together physically, or communication costs in bringing them together by mail, telephone or other means. There may be costs of preparing materials and exhibits for the members of the collectivity, and of paying speakers for them to hear.

In addition to resource costs, there are psycho-social costs of making collective decisions. For example, if an academic department tries to reach a consensus on how money available for raises ought to be allocated among them, they may experience a deterioration in goodwill among themselves that is a cost of that way of making collective decisions.

When collective decisions are made by extortion, additional procedural costs are encountered. These are the costs from deliberate destruction of resources,

injury to persons, pain, loss of freedom and death. There are also psychological costs from the threat of these events. A final and very significant element of procedural cost from extortion is the expense that individuals incur to avoid being subjected to the direct costs of extortion.

The combination of outcome efficiency and procedural efficiency would fully account for the attractiveness of collective decision procedures only if individuals were entirely selfish, and if they believed either that money always had the same value to all persons or that it was not possible to identify factors that are correlated with the value of money. But these conditions are not met. People are not entirely selfish. And it is generally unreasonable to suppose either that money always has the same value to all persons or that it is not possible to identify factors that are correlated with its value. Thus the distributional considerations of equity and stability, as well as efficiency, enter into the evaluation of collective decision procedures.

Equity

Equity is a distribution of goods that conforms to shared understandings regarding what people ought to receive. The taxonomy identifies three aspects of equity: symbolic equality, material equality, and productivity-based inequality. Symbolic equality is the equality of one-person-one-vote, of equality before the law, of equality of opportunity. Symbolic equality possesses the attraction of the idea that no one of us is inherently more worthy than any other. Material equality, on the other hand, is equality in the control of physical resources. Material equality is measured, inversely, by an index of inequality like the Gini coefficient,[1] applied to income or wealth. The attraction of material equality comes substantially from the utilitarian idea that additional resources are more valuable to those who have little than to those who have much. Productivity-based inequality is allocation according to the value of what each person adds to the output of a collectivity by his or her efforts (technically, the person's marginal product). The attraction of productivity-based inequality is illustrated by the widespread agreement that those who gratify their inclinations toward laziness ought to receive less than otherwise similar persons who overcome such inclinations.

There are four interrelated but distinct reasons why it is reasonable to expect concern for the two aspects of equality to influence preferences for collective decision procedures. First, many persons find symbolic equality an attractive ideal, something they strive to express in their private and public lives. Promoting material equality is a way of giving expression to a commitment to symbolic equality.

Second, the implementation of collective decisions often requires individuals to act against their short-term material interests. One way of inducing individuals to

do so is to inspire loyalty in them, and commitment to symbolic equality among the members of a collectivity is instrumental in generating loyalty.

Third, the level of one's material resources is the most obvious factor that is likely to be correlated (negatively) with the value of money. By choosing a collective decision procedure that sacrifices some efficiency for equality, one is, in effect, purchasing wealth insurance: One's expected wealth is lower, but in the event that one is among those who do not fare so well, more assistance is provided.

Fourth and finally, people tend to care differentially about others and to have more concern for those with less wealth than for those with more. Therefore many people value changes that increase the degree of equality among others than themselves. By choosing collective decision procedures that generate more equality, individuals advance their interest in securing greater equality among others than themselves.

While much of our concern for equity is captured by the concepts of symbolic and material equality, productivity-based inequality attracts us even though it diminishes equality. One of the most widely-quoted statements of the virtue of productivity-based inequality occurs in John Locke's *Second Treatise of Government*:

> The labor of his body, and the work of his hands, we may say, are properly his. Whatsoever then he removes out of the state that nature has provided, and left it in, he has mixed his labor with, and joined to it something that is his own, and thereby makes it his property. It being by him removed from the common state nature placed it in, has by this labor something annexed to it, that excludes the common right of other men. For this labor being the unquestionable property of the laborer, no man but he can have a right to what that is once joined to, at least where there is enough, and as good left in common for others.[2]

That productivity-based inequality is attractive is easiest to see at the extremes of productivity. Consider people who are bad-tempered and compulsive liars. It is possible that if we were convinced that these qualities were inherited, we would want to compensate individuals who possess them for the misfortune of their genes. However, we believe that there is at least an element of free will in individual behavior. And we want individuals to receive less when, through their free will, they diminish the well-being of others by expressing their tendencies toward bad temper and compulsive lying than when they control these tendencies. Thus complete equality is not what we seek.

At the other end of the productivity spectrum, consider individuals like Shakespeare, Mozart, Jefferson and Einstein. We acknowledge our great debt to their efforts and are pleased when they achieve more than average material rewards. Even if their contributions rolled from them without great effort, it seems unlikely that our impetus toward equality would be so unwavering that we would seek to limit these highly productive individuals to shares no larger than those of

others. Thus our understandings of what individuals ought to receive incorporate productivity-based inequality.

While the attraction of productivity-based inequality is easiest to discern at the extremes of productivity, the idea is attractive throughout the range of productivity. Thus whenever a person's freely-chosen actions result in more being produced, it is attractive that the producer have discretion over how the additional value is allocated.

Stability

Any pattern of social coordination necessarily includes elements of stability. At a minimum, it is essential that almost all words mean today what they meant yesterday. From an economic perspective, the desire for stability in collective decisions can be characterized as a desire to avoid outcomes that reduce the well-being of any member of the collectivity. If complete stability were attained, then every collective decision would be Pareto improvement, a change that made at least one person better off and no one worse off.

One important attraction of stability comes from its contribution to procedural efficiency. The issue of who is to control what can be debated and fought over endlessly. Much of this conflict takes the form of some persons striving to seize control over things from others, while those who would be dispossessed by these proposals resist. (There are also instances of individuals seeking redistribution from which they would receive no direct personal benefits, but these are less prevalent.) The resources devoted to striving to secure what others have and to striving to retain things currently possessed usually represent waste from the perspective of the collectivity as a whole. This is particularly true when the efforts of the two groups exactly offset each other, so that no net change in allocation occurs. Thus it is attractive to have collective decision procedures that foreclose such wastes of resources.

A related attraction of stability arises from the understanding that it should be possible, in principle, to introduce every improvement in efficiency as a Pareto improvement. That is, for every option that makes some persons better off and others worse off but is on net advantageous, there are in principle modified options differing only in financial compensation from those who would benefit to those who would lose, making everyone better off. The identification of such modifications of proposals that appear to be on net advantageous is highly problematic in practice, but the conviction that they must exist in principle adds to the attraction of distributional stability.

Another attraction of stability arises from the "ratchet effect" with respect to wealth changes. People generally regard the increment in their well-being that comes from a given increase in their wealth, say from W_1 to W_2, as smaller in magnitude than the decrement in their well-being that results from a decrease in

their wealth from W_2 to W_1. Thus a rearrangement of the identities of the individuals who receive different incomes, leaving the distribution of incomes unchanged, generally produces an overall reduction in well-being.

While stability is thus attractive, in many situations there are conflicting conceptions of what stability entails. This is reflected in the second level of the taxonomy, where stability is divided into "restoration" and "preservation." Restoration is the aspect of stability entailed in a commitment to undo any illegitimate departure from some initial status quo. Preservation is the aspect of stability entailed in a commitment to resist any departures from the current status quo. Preservation can conflict with restoration because preservation incorporates a commitment to resist even those departures from the current status quo that would restore a previous status quo.

The appeal of restoration is clearest in the shared understanding that things that have been stolen must be returned to those from whom they were stolen. The conflicting appeal of preservation is revealed by the diminution, as time passes, in the impulse to undo the consequences of theft and other illegitimate changes. Thus there is a statute of limitations for prosecuting thieves. The passage of time also tends to diminish the appeal of restoration by making it difficult to identify what restoration entails. If my father stole a painting from your father and passed it on to me, I could be required to restore the painting to you, but if my father sold the painting and used the money to buy me an education that I might not have chosen for myself, what does restoration then entail?

That the passage of time does not invariably extinguish the salience of restoration may be seen in riots that occurred in India in 1990. Hindus claimed that 300 years earlier, when Moslems controlled India, a mosque was built on a site where a Hindu temple had previously stood. They wanted the mosque torn down so that the temple could be restored.

While it may seem extreme to attempt to rectify a 300-year-old injustice, it is also troubling to allow the mere passage of time to legitimate a new status quo. In the rhetoric surrounding Iraq's take-over of Kuwait in 1990, one saw efforts by Iraqi leaders to give the new situation the imprimatur of the status quo, while other world leaders strove to perpetuate a perception of the impropriety of Iraq's actions and the need to restore the previous status quo. World opinion reflects an awareness that if the imprimatur of the status quo is easily attained after a coercive change, then aggression will be all the more attractive and prevalent. And so the tension between preservation and restoration persists.

Conclusion

Each of the general criteria of efficiency, equity and stability has a rationale that seems to allow no compromise with the other criteria. If efficiency is attractive because collective decisions ought to provide as much as possible of whatever

people want, then what place is there for any other criterion? But similarly, if equality and productivity-based inequality are attractive because we have shared understandings about what individuals ought to receive, how can any other criterion be relevant? And finally, if our concern for stability reflects a determination not to allow change to make any individuals worse off, how can we compromise on that position without making it meaningless?

It appears that the attempt to identify criteria for evaluating collective decision procedures has the unintended consequence of creating slogans that tend to exclude the attraction of competing criteria. But the fact that the competing criteria are attractive when we focus on them reveals that each slogan overstates the attraction of the general criterion for which it speaks.

To understand what we really want from collective decision procedures we must not succumb to the monistic tendencies suggested by the rationales for individual criteria. We must be alert to the attraction of criteria in all categories as we ask how the set of consequences of each collective decision mode compares with those of all other modes.

The next chapter provides a discussion of the relative advantages of the various modes in terms of the criteria identified in this chapter.

Notes

[1] The Gini coefficient can be defined as the sum over all pairs of individuals of the differences in their shares of some total, divided by (the number of individuals – 1). It has a range of 0 to 1.

[2] John Locke, *Two Treatises of Government* (New York, 1960), Second Treatise, section 27, pp. 328–9.

This page intentionally left blank

Chapter 5

Relative Advantages of Modes of Making Collective Decisions

Introduction

This chapter evaluates modes of making collective decisions in terms of the criteria identified in Chapter 4. The evaluations are summarized in Table 5.1 on the next page.

The scores in Table 5.1 are impressionistic in the sense that, while explanations can be given for them, they cannot be defended as unquestionable. Nevertheless, the scores serve to organize thinking about which modes are attractive, and why. A score consisting of one or more pluses indicates that a mode is attractive in terms of a criterion; the more pluses the better. Similarly, one or more minuses indicates a notable failure in terms of a criterion; the more minuses the worse. A combination of one or more pluses with one or more minuses indicates a situation in which performance has a notable range of possibilities. A zero indicates a neutral performance in terms of a criterion.

In the next section the criteria are considered individually. As each criterion is considered, modes are evaluated in terms of it. In these subsections, modes are taken up in approximate order of their attractiveness in terms of the criterion under consideration. In the following section the modes are considered individually, their strengths and weaknesses discussed, and the circumstances in which they are most likely to be attractive are identified. The final section summarizes the conditions leading to choice of various modes.

Criteria

Procedural Efficiency

Authority (when the authority is a single individual) and random process are the modes that rank highest in terms of procedural efficiency. Procedural costs for these modes are minimal. Voting is next most attractive. Voting involves costs of informing voters, of going to the polls, and of processing the ballots. (Some of

Table 5.1 The attractiveness of modes in terms of criteria

	Efficiency		Equity			Stability	
	Procedural efficiency	Outcome efficiency	Symbolic equality	Material equality	Productivity-based inequality	Restoration	Preservation
Consensus	+ + + – – –	+ + +	+ + +	+ + +	+	+	+ + +
Pseudo-consensus	+ + + – –	+	+	+ +	+	+ +	+
Trade	+ + + – –	+ +	+ +	0	+ + +	+	+ +
Extortion	+ – – –	0	– –	–	– –	+	– –
Authority	+ + +	+	–	+ –	+	+ + +	+ –
Random process	+ + +	–	+ + +	–	0	0	–
Voting	+ +	+	+ + +	+	+	+ +	–
Contest	+ –	+ + –	+	–	+ +	+	0

these costs do not arise when voting is used in more informal settings.) As high as these costs can be on some occasions, they are almost always lower than the costs of using any of the modes in the agreement-on-outcome family.

The procedural efficiency of contest is highly variable. It tends to be greatest when there are no expenditures that the advocates of an option can make to improve its prospects of being adopted. Under some circumstances (as when the contest is a court case) the expenditures of the advocates of different options substantially offset each other, leaving the aggregate of the expenditures as a cost that yields no net social benefit. On the other hand, there may be instances (referenda where proponents and opponents make substantial expenditures) where the expenditures of advocates represent an investment that is essential to ascertaining the relative value of the options.

All of the modes in the agreement-on-outcome family have procedural efficiency that is highly variable, and all of them typically have lower procedural efficiency than the modes in the agreement-on-procedure family. However, it is also true that in those instances when they are most efficient (a purchase or an immediate consensus between two persons), the modes in the agreement-on-outcome family are as efficient as any mode in the agreement-on-procedure family. A further generalization that can be made about the modes in the agreement-on-outcome family is that their procedural efficiency is extremely sensitive to the number of persons in a collectivity. These procedures have the greatest procedural efficiency when the number of persons in a collectivity is smallest.

Among modes in the agreement-on-outcome family, the highest procedural efficiency is generally possessed by trade when the number of persons involved is small, and by pseudo-consensus when the number of persons involved is large. With trade, the possibility of offering compensation overcomes what might otherwise be insurmountable differences regarding what ought to be done. On the other hand, the fact that compensation is possible can lead to "hold-out" behavior that reduces procedural efficiency. The procedural efficiency of trade depends not only on the number of persons who must agree to implement a trade, but also on the number of alternative groupings that can substitute for any particular collectivity. Thus a substantial part of the efficiency of competition arises from the multiplicity of alternative possible sets of trading partners. The procedural efficiency of trade falls considerably when there are no substitutes of trading pairings (bilateral monopoly) and even more so when (as in land assembly) there is a large number of specific persons who must agree to implement a trade. The knowledge that their consent is required leads some persons to insist that they receive disproportionately large shares of the gains to be made from a trade. The difficulty of satisfying demands of this sort can make the procedural costs of trade very high and even prohibitive.

Pseudo-consensus substitutes pressure on members of a collectivity to agree, in place of the compensation that is offered when trade is used. This generally

achieves results superior to those of trade in large-number situations because it provides no opportunity for hold-out behavior.

The procedural efficiency of consensus can be very high when basic values and understandings are shared in a collectivity. When outlooks are sufficiently different, consensus can be impossible to achieve, in effect giving consensus an abysmally low procedural efficiency.

The procedural efficiency of extortion tends to be lowest of all. All of the problems of trade apply to extortion, and these are magnified by the lack of agreement regarding what is proper conduct in the absence of agreement, and by hostility among the members of collectivities. All the destruction of wars, the costs of armies, most of the expenditures on police forces, and the expenditures on arms by criminals and by those defending themselves from criminals are components of the procedural inefficiency of extortion.

Outcome Efficiency

When it is possible to achieve consensus, the mode of consensus is the most attractive one in terms of outcome efficiency. If all the members of a collectivity are able to agree that an outcome is right, then that outcome can be presumed to be highly efficient. The outcome efficiency of consensus is not perfect because not all possible outcomes come to the attention of collectivities, and the evaluation of those that do come to their attention is sometimes flawed. When it is not possible to achieve consensus, the effort to use consensus can cause significant losses in outcome efficiency. Either there is no collective decision, and the benefits that could have been realized from coordination of individual actions are forgone, or the mode of the collective decision making process is transformed, consciously or unconsciously, from consensus to pseudo-consensus, trade or extortion.

Trade is second best in terms of outcome efficiency. In addition to the outcome inefficiency arising from the sources that affect consensus, trade incurs additional outcome inefficiency from the effort to reduce losses from procedural inefficiency. People will agree to an outcome that may be less than the best possible trade for the sake of insuring that some agreement is reached. The outcome efficiency of trade is also limited by the problem of making commitments credible, which does not arise with consensus. There are three ways in which the commitments entailed in trade can achieve credibility. First, trade can occur in the context of a consensus within a collectivity that commitments are kept. Because of such a consensus, individuals keep their own commitments and trust that others will keep theirs. Second, individuals may understand the value of acquiring a reputation for keeping commitments. Those who stand to profit from having such reputations in the future will be trusted to behave in ways that will protect their reputations. Third, a judicial system can be used to give individuals material incentives to keep their commitments. Through the combination of these devices, trading commitments

are given enough credibility for trade to be reasonably outcome-efficient in many circumstances.

The outcome efficiency of contest varies in a way that is correlated with its procedural efficiency. The outcome efficiency of contest is high when, as in competitive bidding, it is possible, at relatively low cost, to conduct a contest that can be relied upon to identify the best option irrespective of expenditures by the advocates of options. The outcome efficiency of contest is lower when the result of the contest is readily affected by expenditures of advocates.

Pseudo-consensus gains a measure of outcome efficiency from the fact that all individuals in the collectivity must consent to the outcome of a collective decision. On the other hand, the fact that no compensation can be provided will often cause individuals to find it impossible to consent to outcomes that would be efficient. Thus pseudo-consensus is not very high in outcome efficiency.

The claim of authority to outcome efficiency has two bases. First, the person or persons in authority may be wiser than the average member of the collectivity. If the authority is both wiser than average and benevolent, then the outcomes selected by authority can for that reason be superior to those selected by any other mode. The second basis for the claim of outcome efficiency by authority is based on the claim that authority has high procedural efficiency. In a large collectivity, a single individual can generally expect to have little if any influence on outcomes. This reduces the incentive of any individual to become informed about the relative merits of the available options. And an uninformed collectivity is less likely to select the option with the greatest value. By concentrating influence on a small number of individuals or on a single one, the mode of authority increases the incentive for those with influence to become informed, which improves outcome efficiency.

The primary claim of voting to outcome efficiency rests on a conception of symmetry of interests.[1] If the average intensity of the interest of those in the majority is the same as for those in the minority, then outcome efficiency lies with following the wishes of the majority. However, because the average intensity of interests of the majority and the minority are not always equal, voting does not achieve perfect outcome efficiency. Nevertheless, the tendency toward equality of interests makes voting attractive in many circumstances. Furthermore, there are varieties of voting that, to some extent, overcome reliance upon an assumption of symmetry of interests and therefore achieve greater outcome efficiency. These are discussed in Chapter 16. Another important limitation on the outcome efficiency of voting is the lack of incentive for members of a collectivity to become informed, in view of the very low probability that the vote of a given individual will be decisive.

Extortion has a low outcome efficiency for two reasons. First, the procedural efficiency of extortion is low, so it loses substantial outcome efficiency from the effort to save procedural costs. In addition, it suffers from a limited ability to make commitments credible. The parties to a potential extortion are unlikely to be

members of a collectivity sharing a consensus that commitments are kept. And they are unlikely to be subject to a court system that will be able to impose sanctions on them. (There are exceptions: The participants in labor disputes, for example, may regard as obligatory the behavior that they are seeking to persuade others to agree to—thereby making the collective decision mode extortion—while recognizing that it is not *legally* obligatory unless it is incorporated into a contract.) Reputation, the third mechanism for giving commitments credibility, does operate under extortion. For example, Alexander the Great was aided in his conquests by a reputation for keeping his promises to those cities that submitted to him without a fight.

Random process has virtually no claim to outcome efficiency. If there is a claim of random process to outcome efficiency, it rests on the idea that "random processes" are not actually random, and can be used as a channel for superhuman guidance. This idea appears to lie, for example, behind the Biblical accounts of the selection of Saul as the first king of the Israelites (I Samuel) and the selection of Matthias to replace Judas Iscariot among the Apostles (Acts).

Symbolic Equality

There are three modes that are equally excellent in terms of symbolic equality: consensus, random process, and voting. In consensus, the symbolic equality comes from the fact that all must agree, in random process it comes from the fact that all have equal chances, and in voting it comes from the fact that each person has one vote. (There are random process and voting rules without these symmetries, but the important point is that if symbolic equality is desired these modes can provide it.) Random process is particularly useful for achieving *ex ante* symbolic equality when it appears to be impossible to achieve *ex post* equality. One team has to make the opening kick off; the other gets to receive. Flipping a coin for the kick-off gives the two sides equal chances to be favored. If the army is raised by a draft, someone has to be drafted. Selecting the draftees by a lottery gives everyone an equal chance to bear the burden.

The symbolic equality of trade lies in the fact that each participant in the trade must be satisfied for the trade to occur. The limitation in the symbolic equality of trade lies in the possibility that the endowments from which trading begins may be so unequal that no symbolic equality in trade can be perceived.

Pseudo-consensus retains some of the symbolic equality of consensus, from the fact that all must give their assent, but the fact that individuals are pressured and do not actually agree dilutes its symbolic equality.

The symbolic equality of contest lies in the fact that the rules can be designed so as to give no individual a privileged status. The limitation on the symbolic equality of contest comes from the possibility that even though the rules are even-handed, the natural endowments of individuals are such that some are much more likely to win than others.

The asymmetry involved in the award of authority tends to deprive it of any symbolic equality. Nevertheless, authority has the capacity for a trace of symbolic equality, from the possibility of even-handed decisions by wise, benevolent, disinterested persons.

The only symbolic equality of extortion is in the symmetry of a Hobbsian "war of all against all."

Material Equality

If it is possible to achieve consensus, then whatever level of material equality a collectivity desires is achieved. However, if it is not possible to achieve consensus, then the level of material equality is determined by the lack of a collective decision or by the characteristics of whatever mode is used instead. If the mode is pseudo-consensus, then it is likely that the need to secure the consent of all will result in a substantial amount of material equality, if that is generally desired.

Voting is highly variable in the amount of material equality it produces. The fact that all votes count the same has a tendency to generate outcomes with substantial material equality, although this is offset to some extent by two interrelated factors. The first factor is that when voting is used, a majority coalition is able to obtain disproportionate benefits for itself at the expense of those outside the coalition, reducing material equality. If all coalitions had the same chance of forming, this would average out in the long run. But coalitions do not all have the same chance of forming. The second factor is that individuals are more likely to be in the majority coalition if they share with others characteristics that promote loyalty to a coalition, or if they are better at the game of politics. Thus voting, to some extent, promotes material inequality.

Authority is mixed in its ability to achieve material equality. The positive contribution of authority to material equality comes from the possibility of wise, benevolent, disinterested judgment by those in authority, giving a collectivity the amount of material equality that is widely seen to be appropriate. The negative aspect of authority with respect to material equality arises when those in authority determine how much they themselves will receive and use this power to award themselves more than they deserve.

Trade does little to promote or diminish equality within a collectivity. There is a tendency for bargaining to yield outcomes that split the gains from trade roughly equally, but there is also substantial dispersion around this central tendency. Furthermore, whatever inequalities exist prior to trade tend to be perpetuated after trade.

Collective decisions made by extortion have haphazard effects on equality. The contributions to equality come from such things as the poor robbing or kidnapping the rich or receiving largess from "Robin Hoods" who rob the rich. However, the poor tend disproportionately to be victims of robbers—they tend to

lack the resources that the rich have to defend themselves. Furthermore, the willingness to engage in the extortionate behavior that leads to personal gains is not uniformly distributed. The gains from collective decisions made by extortion are concentrated on those who do not mind trampling on what others see as their rights.

Decisions made by random process and contest tend to diminish the material equality of a collectivity. With both modes, some win and others lose. With random process, all have equal chances to win or lose, while with contest the better endowed are disproportionately likely to win.

Productivity-Based Inequality

The mode of trade is superb for identifying productive individuals and rewarding them according to their productivity. If one trading partner is unwilling to pay a person an amount commensurate with what he or she produces, then another trading partner who is willing to do so can be found. The mode of contest is fairly good for securing productivity-based inequality, although it has the potential to focus excessively on aspects of productivity that are readily susceptible to measurement. For example, if a competition is held for a position in an orchestra, there is likely to be more attention to the ability of applicants to play solos than to their ability to blend their playing with that of the other members of the orchestra. Consensus, pseudo-consensus, authority and voting can all be expected to have some success in achieving productivity-based inequality, since those who participate in decisions have some tendency to reward productive individuals, to give them incentives to remain productive. Random process contributes nothing to productivity-based inequality, and extortion actively diminishes it.

Restoration

Restoration requires active intervention to regain a previously established status quo. When an effort is made to achieve restoration, there are two types of collectivities that might be analyzed: a collectivity composed of all parties who would be involved in the restoration (a thief, a victim and a mediator), and a collectivity capable of imposing restoration on a dissenting initiator of a previous change (a victim and representatives of the judicial process imposing restoration on a thief).

Restoration is most satisfying when the initiator of an earlier change is included in the collectivity that implements the restoration and the mode that is used is consensus. This, however, is rarely achieved. The necessity of undoing a change that has previously been implemented, usually by someone who thought it would endure, makes restoration against the wishes of an initiator of change difficult and sometimes impossible to attain. The mass mobilization that is needed to impose restoration on a dissenting initiator is most readily generated by the mode of

authority. The modes of pseudo-consensus and voting also have a notable capability for this task, although a selfish collectivity may ignore such appeals.

Rewards for information leading to the return of stolen property or for the capture of persons who have jumped bail represent the use of contest (if anyone can collect the reward) or trade (if a particular person is hired) to achieve restoration. Paying a reward to a thief for the return of property has some of the qualities of trade, but more of the qualities of extortion, since the victim of the theft would generally assert that the thief ought to return the stolen property without the incentive of payment. The fact that a payment is made to the thief means that only partial restoration is achieved. Random process contributes virtually nothing to restoration.

Preservation

When it is possible to achieve consensus, the mode of consensus provides as much preservation as a collectivity desires. Consensus does not guarantee complete preservation, because sometimes a consensus will be reached on an outcome in which some members of a collectivity agree to contribute without receiving corresponding benefits. Such events are not indications of limitations of consensus, but rather of the fact that collectivities do not always want preservation.

The mode of trade provides a high degree of preservation. Any trade makes every member of the collectivity involved in the trade at least as well off as he or she would be in the absence of the trade. If there is a difficulty with trade in terms of preservation, it is that there is sometimes disagreement about the rights people have before trade begins. If a firm decides that in the future it will offer lower bonuses for overtime, is this an infringement on the rights of employees to their traditional compensation, or simply a difference between the terms of past trades and the terms of future trades? People tend to develop expectations about the terms of future trades. They often disagree about whether there is an obligation to continue past terms of trade in the future. In these circumstances, what appears to one person to be an offer to trade on terms that happen to differ from past terms, appears to another person to be an abrogation of an obligation and an introduction of distributional instability through trade. Thus the ability of trade to achieve preservation is crucially dependent on a common understanding of rights and obligations prior to trade.

Pseudo-consensus is reasonably good in terms of preservation. Losses to individuals can occur, but these are limited to what individuals can be persuaded to accept.

When those who are wise and benevolent exercise authority, the results can be reasonably good in terms of preservation. Authority is particularly attractive as a way of achieving approximate preservation in such collective decisions as the amount of compensation to be paid for property assembled by eminent domain, where the number of participants tends to make trade impractical. Wisdom,

however, is never able to insure that collective decisions imposed by authority actually achieve preservation. Authority exercised unwisely or without benevolence can do extremely poorly in terms of preservation.

The impact on preservation of collective decisions made by contest tends to be haphazard except when the impact on preservation is susceptible to measurement and is employed as a criterion of the contest by which the outcome is selected. For example, a competition for plans to redevelop a neighborhood might be judged in part by the quality of the relocation services offered to the residents of the neighborhood.

Voting tends to be not very good in terms of preservation, because it permits a majority to impose its will on the minority. The damage that voting does to preservation is limited to some extent by constitutions, and also by the tendency for logrolling and other forms of vote trading to emerge, so that every member of a collectivity is likely to gain from some collective decisions made by voting, even if he or she loses from others.

Random process is about on par with voting in terms of preservation. No individual collective decision made by random process can be expected to provide preservation. For a series of collective decisions, however, the tendency of random effects to cancel one another serves in place of vote trading to reduce the likelihood of very large losses for individuals.

Because a collective decision by extortion entails the *quid pro quo* of trade, it has an appearance of providing preservation. However, this appearance is relative to a starting point that incorporates any threats and predations that precede the collective decision. It is these threats and predations that generate the distributional instability of extortion. Because there is little to limit their magnitude, extortion is the least attractive mode in terms of preservation.

Modes

Consensus

Consensus is the mode for making collective decisions to which collectivities generally aspire. The appeal of consensus is connected to the belief that there is a unique appropriate outcome of a collective decision in any set of circumstances. If only, it is thought, the entire collectivity could avoid erroneous thinking and perceive the issue properly, there would be consensus upon the appropriate outcome.

Whenever consensus can be discovered or achieved, decisions made in this mode provide outcome efficiency, material equality, symbolic equality and preservation that cannot be surpassed by any other mode. The “only” problem with consensus is that there is no way of insuring that consensus will be achieved, and the effort to achieve it can be time-consuming, costly, and frustrating. Thus

collectivities tend to seek to make their collective decisions by consensus when and only when they expect to share a single outlook.

For the goals of restoration and productivity-based inequality, which require complex action, the low procedural efficiency of consensus can be very costly. Frustration with the difficulty of achieving consensus tends to lead to a transformation of the mode from consensus to pseudo-consensus, trade, or extortion, sometimes unconsciously.

Pseudo-Consensus

Pseudo-consensus is a variation on (or perhaps a transformation of) consensus that accommodates the fact that sometimes it seems impossible to achieve complete agreement. Pseudo-consensus is not as susceptible as consensus to extremely poor procedural efficiency, but it does notably worse than consensus by most other criteria. Its outcome efficiency is good but not outstanding, because there is no way to guarantee that the right voices will be followed when some are ignored. Its material equality is reasonably good because only outcomes with the simplicity of equality have a chance to attract a near-consensus. Its symbolic equality suffers because of the disadvantaged position of those who merely tolerate the outcome. Pseudo-consensus does reasonably well in terms of productivity-based inequality, restoration and preservation because outcomes that promote these objectives will be favored by the preponderance of members of collectivities.

Trade

Trade is a variation on consensus that acquiesces to the selfishness in people. Instead of asking what is right, trade asks what can command agreement. By offering compensation for participation in an outcome, trade is able to work in situations where consensus cannot be achieved. But providing compensation has a drawback. Trade invites people to ask whether they might be able to secure even more compensation by withholding their participation until more compensation is offered. When numerous different groupings for trade are possible and the number of persons involved in each trade is small, this drawback is inconsequential. When only a single grouping is possible but the number of persons is small, the selfishness invited by trade still does not usually thwart collective decisions. When the number of persons is large, the procedural efficiency of trade tends to fall to unacceptable levels, though this is less true when a variety of alternative groupings are possible.

When agreement can be achieved, the outcome efficiency of trade tends to be quite good. Trade does extremely well in terms of productivity-based inequality, and is also quite good in terms of preservation, though it must be noted that to keep trade from being transformed into extortion with its low preservation, there must be agreement about what obligations prevail.

Trade offers the symbolic equality of equal opportunity to engage in trades and contributes to restoration through private contracts to recover property. The one criterion by which trade necessarily scores low is material equality.

Extortion

Extortion is the mode of collectivities that do not have enough cohesion to make collective decisions by other modes. It has nothing to recommend it as a deliberately chosen mode for making collective decisions when other modes are feasible. The difficulty of making credible commitments without cohesion has an adverse effect on the outcome efficiency of extortion. The effort to make threats credible gives extortion an appalling procedural efficiency. The effect on equality of collective decisions made by extortion is haphazard. They generally undermine preservation, and only occasionally do they provide restoration.

Authority

The relative merit of authority is greatest with respect to procedural efficiency and restoration. Authority can get things done and discipline wrong-doers. In outcome efficiency and productivity-based inequality, authority is of some value but is limited by its reliance on what one person, or in any case a subset of the collectivity, can grasp. On the other hand, the possibility of giving authority to wise persons provides a possible advantage for authority.

With respect to material equality and preservation, authority can perform reasonably satisfactorily if the authority is not only wise, but also disinterested and yet benevolent. The possibility that authority will fall into the hands of persons who are unwise and/or selfish gives authority the potential for quite poor performance in material equality and preservation. The symbolic equality of authority is inherently low.

Random Process

The greatest virtues of random process are by the criteria of symbolic equality and procedural efficiency. It provides *ex ante* symbolic equality in circumstances when it seems impossible to avoid *ex post* inequality, and it costs almost nothing to implement. However, it performs quite poorly in terms of all other criteria because outcomes incorporate no intelligence.

Voting

Voting is almost as good as authority and random process in terms of procedural efficiency. It tends to be reasonably good in terms of outcome efficiency—not as good as consensus or trade, but often better than authority, with less risk than

authority of poor outcome efficiency. Its orientation to action makes voting useful for restoration.

Voting is excellent in terms of symbolic equality and has some merit in terms of material equality and productivity-based inequality, although its susceptibility to continuing dominance by a selfish majority coalition gives it a risk of performing poorly by these criteria. The regular existence of disappointed minorities gives voting a poor rating in terms of preservation.

Contest

The performance of contest by all criteria is highly variable. Good performance can be expected to the extent that it is possible to achieve agreement on the manner in which outcome efficiency, equity and stability are to be combined in evaluating options, and then design a contest that reveals unambiguously the performance of options in terms of these criteria. Unfortunately, the required consensus about combining the criteria is rarely achieved, and even when it is, performance is rarely unambiguous, so that some other mode is needed to settle disagreements about who won the contest. The possibility of ambiguity in contest outcomes and of altering outcomes through expenditures by advocates create significant difficulties for contest in terms of all the criteria by which it might do well.

Conditions Leading to the Choice of Various Modes

The manner in which the modes vary in performance by the stated criteria leads in a natural way to different choices of mode depending on circumstances. When the members of a collectivity expect to be able to share a single outlook, they will seek to make their collective decisions by consensus. When a common understanding cannot be expected, some other mode must be used if any collective decisions are to be made. If a shared outlook is nearly universal in a large collectivity, then the mode of pseudo-consensus may be attractive. If the number of persons in the collectivity is very small, or if alternative groupings into collectivities are possible, then consensus or trade is likely to be chosen. A decision that is supposed to be made by consensus will sometimes be made by trade or extortion without an explicit decision to change the mode.

A collectivity's first decision to use a mode from the agreement-on-procedure family must be made by a mode from the agreement-on-outcome family. A decision to use this secondary family of procedures is based on a perception that it will be too costly, if not impossible, to reach complete agreement on an outcome. Within the agreement-on-procedure family, the mode of contest has the attraction of potentially regenerating complete or nearly complete agreement. However, if this were always possible, the contest could be imbedded in a process relying on consensus or pseudo-consensus. Contest retains an attraction as a way of

approximating consensus, although it then requires a formal apparatus to determine the winner of the contest. Voting is the mode that usually seems most natural for identifying contest winners. It is particularly attractive when the members of the collectivity are seen as having the same underlying interests, but as having different perceptions of how those interests are best served. Authority (possibly a plural authority using voting to resolve their differences) is attractive when significant effort or unusual ability is required to discern the most appropriate option. Random process is attractive for resolving voting ties (or in two-person collectivities where all votes that were not unanimous would be ties) and for making choices among options that determine who will gain and who will lose when it seems inevitable that there be gainers and losers.

The next chapter addresses the question of what is meant by a "good collective decision."

Note

1 See Howard R. Bowen, "The Interpretation of Voting in the Allocation of Economic Resources," *Quarterly Journal of Economics* 58 (1943): 27–48.

Chapter 6

What is a Good Collective Decision?

Introduction

At the level of abstract theory, a collective decision is simply a coordination of intentions. No normative evaluation of the outcome can be deduced logically from the fact that it is chosen. But even if a logical deduction cannot be made, there is a strong tendency to expect chosen outcomes to be good. Discussions of options tend to be discussions not of, "What shall we do?" but rather of, "What should we do?" To ask, "What should we do?" suggests that the question has a definite answer, and that it ought to be possible to know it. But that may not be possible, and the failure to deal with this can lead to serious misunderstandings. For this reason it is important to explore in some detail how knowing occurs.

This chapter begins with a general discussion of what knowing is and how it is accomplished, as a prelude to a discussion of how knowing applies to what a collectivity ought to do, and why the issue is important for collective decisions. This is followed by a discussion of the way that knowing what a collectivity ought to do might be applied to each of the modes of collective decision-making. The subsequent section discusses criteria for good collective decision rules in general. A final section summarizes the chapter.

What is Knowing, and How is it Accomplished?

A conception of knowing that is found in the philosophical literature is that person *P* knows thing *T* if:

1. *T* is true.
2. *P* believes that *T* is true.
3. *P* has a good reason for believing that *T* is true.[1]

It appears that no one doubts the sufficiency of the three stated conditions. They are regarded as all necessary because, 1) a person, it would appear, cannot know something if it is not true; 2) even if something is true, a person clearly does not know it if he does not believe that it is true; 3) even if something is true and a person believes that it is true, we would not say that he knows it if he does not have

a good reason for his belief. For example, if the New York Yankees won the 1956 World Series, and if I believe that the New York Yankees won the 1956 World Series, most people would not agree that I *know* that the New York Yankees won the 1956 World Series if the reason I give for my belief is that New York and Yankees both have seven letters, and 56 is divisible by seven. To qualify as knowing, a person who believes that something is true must have an acceptable reason for his belief.

This conception of knowing makes it necessary to determine whether a thing is true before it can be determined whether anyone knows it, and this is a significant limitation. It may seem strange to call this a limitation, for how could it possibly be asserted that a person knew a thing if it were not known whether the thing was true? The answer is roundabout.

First it must be noted that changes in general opinion concerning what things are unquestionably true occur in the widest variety of places:

1. In the 1960s at least one young geologist was denied tenure for supporting the then-heretical view that the earth's crust was composed of plates that moved about and slid under each other. That view is now orthodox.

2. It is not clear that anyone bothered to ask the question, but if asked whether the mass of an object was independent of its velocity, it is hard to imagine any nineteenth century physicist pausing more than a second before responding, "Of course." Modern theories of relativity say that it is not.

3. For centuries mathematicians understood that negative numbers did not have square roots. Then "imaginary" numbers came along. The question of whether negative numbers have square roots is now understood to be entirely a question of the framework in which the question is asked.

4. For centuries the dominant view was that it was possible for one human being to own another. Now almost no one makes such an assertion.

5. In the sixteenth century the Church asserted that the earth was the center of the universe and felt so strongly about this that Galileo was threatened with torture for arguing otherwise. It has been a very long time since any theologian has argued that an earth-centered universe was essential.

While I have refrained in these examples from saying that the individuals involved "knew" the things that later came to be seen as not true, it is not clear that the individuals themselves would have refrained. In each case they might well have asserted that they were as sure of the thing in question as they were of anything. And yet it came to be seen to be otherwise.

One possible response to this state of affairs is to say that there is an unfortunate tendency of people to think they know things when they don't, and an appreciation of Bayes' Theorem would help them realize this.[2] This application of conditional probability can be stated as follows: If there are two possible states of the world, A and B, and if, from some unspecified source, the probabilities of the two states, prior to the incorporation of the implications of data D, are P_A and P_B respectively, and if the probabilities of observing data D given the two states are

$P_D \mid A$ and $P_D \mid B$ respectively, then the implication of observing data D is that the probabilities of the two states of the world become

$$\frac{P_A \bullet P_D \mid A}{P_A \bullet P_D \mid A + P_B \bullet P_D \mid B} \quad \text{and} \quad \frac{P_B \bullet P_D \mid B}{P_A \bullet P_D \mid A + P_B \bullet P_D \mid B}$$

respectively.

One important idea that is highlighted by Bayes' Theorem is that mere improbability can never yield the implication that a theory is false. Whenever there are two (or more) explanations that are not logically ruled out by the available evidence, the fact that one theory accounts for the evidence much more parsimoniously than any other can never be sufficient to establish that the theory in question is right and all others are wrong.

From a Bayesian perspective, anyone who says that he knows anything is deluding himself. Evidence is inevitably finite, while the number of possible explanations is infinite. There can never be sufficient evidence to justify an assertion that *all* alternative hypotheses have been refuted. From this perspective, the only interpretation that can be placed on the phrase "I know" is that it is a sloppy approximation for "I believe that it is highly likely that."

The difficulty with this Bayesian interpretation of "I know" is that it does not fit the way people use language. Consider another example. Almost everybody knows that George Washington was the first President of the United States. Nobody says, "I believe that it is highly likely that George Washington was the first President of the United States." They say, "I know that George Washington was the first President of the United States." Wouldn't you? Actually, under the Articles of Confederation, several individuals held the title of President of the Unites States.[3] The first of these, John Hanson of Maryland, not George Washington, was the first person to hold the title of President of the United States.

The purpose of the example is to illustrate that the phrase "I know" can be used in a perfectly understandable way about something that is not true. If the meaning of "I know" is to be found in an inquiry into how it is used, then it can be seen that people do not in general operate in terms of the intermediate probabilities required by Bayes' Theorem, and it is not necessary for the predicate of "I know" to be true for the phrase to be used in a perfectly understandable way.

This illustrates that "know" has two distinct meanings, depending on whether or not one is caught up in the illusion of certainty. From the perspective of a person who says "I know," or of someone who shares the same framework, something that is "known" is true. However, from the perspective of someone who is engaged in an examination of language and beliefs, that which is "known" need not be true. From the latter perspective, *To know is to make a practice of not doubting, and have a reason for that practice that is accepted by the persons with whom one communicates*. In other words, from this perspective the second and

third of the requirements for a thing to be known must be satisfied, but the first need not be satisfied.

Knowing in this sense is a cognitive strategy. It works because human minds and experiences are sufficiently similar. Furthermore, people are sociable and to a substantial degree are capable of being persuaded to adopt the assumptions employed by others. This permits the development of communicating groups that make a practice of not doubting the same things.

We cannot communicate with one another without making assumptions: We assume, without any direct evidence, that others are conscious beings. (A society of solipsists is a contradiction in terms.) We assume that others will understand at least some of the words or signs that we plan to use in the same way that we understand them. Thus no discussion can begin at the beginning. Efforts to communicate involve beginning in the middle and risking incomprehension, which can sometimes be detected.

Consider then the idea that knowing is accomplished by joining others in what they assume or by persuading others to join you in what you assume. This idea feels uncomfortable. We want what is "known" to be "objectively" true, true whether anyone knows it or not. This drive for objective truth in what is "known" may have an evolutionary rationale. It is hard to bring all of one's efforts to bear on a task when one has doubts. Doubts are sometimes productive, but evolution would select individuals that are capable of setting aside their doubts, after entertaining them, and acting with conviction. A person does very little harm to himself by indulging in the delusion that the mass of an object is independent of its velocity, or that George Washington was the first President of the United States, whereas his probability of surviving could be seriously jeopardized by a failure to make logically invalid inferences based on observed regularities. People might starve if they insisted on not eating until it could be incontrovertibly proved to them that their eating did not violate the rights of any other plant or animal to live. (A "proof" establishes a correspondence between a set of assumptions and a conclusion, and a basis can always be found, if one wishes, for not accepting an assumption.)

We strongly desire to have an irrefutable basis for asserting that what we believe is *really* true. But the practical significance of whether we are able to achieve this is slight, and more political than philosophical. There is no general algorithm for determining the truth status of propositions. Things are agreed to be true by a variety of different consensus-creating procedures. Sometimes a proposition is encountered that nearly everyone agrees is true. (The earth is round.) If we encounter someone who says that such a thing is not true, we decide that he must be joking, or else he is crazy. We dismiss such a person, deciding that there is no need to pay attention to what he says on this issue, and quite possibly on other issues as well.

The importance of knowing what is definitely true is connected to concerns about the appropriateness of dismissing people who say things that seem

nonsensical. If we can be absolutely sure that they are wrong, then we can know that we are right to dismiss them. However, if the processes we use to determine what is true are not infallible, then those whom we would like to dismiss may have reasonable unsatisfied claims on us. The question of what we know and how we know it is thus a question of what obligations we have to those who disagree with us. To persist in asserting that we have absolute knowledge is to erect a barrier against those who see things differently than we do. To acknowledge that our ways of knowing are fallible is to become more susceptible to the claims of those who challenge our ideas.

We fight a deeply ingrained instinct when we try to maintain a skeptical open-mindedness about all ideas. In fact, while it is possible to be so hesitant that we are unable to act, we cannot doubt all ideas simultaneously because an idea must be formulated in words before it can be doubted, and this involves us unintentionally in the assumption that we know what the words describing the idea mean. We lack a valid basis for asserting that what we do not doubt is incontrovertibly true, but we are unable to doubt all that is dubious. All we can do is to nibble at the edges of our convictions, here and there replacing an identifiably dubious assumption with an assumption whose limitations are not yet visible. But we get by nevertheless, leading lives whose value many of us are able to refrain from doubting.

Ethical Knowledge

Ethical knowledge has the same structure as factual knowledge. Like-minded groups join together in not doubting factual postulates and in not doubting ethical postulates. Freudians do not doubt that a person has an id, an ego and a super-ego. Chemists do not doubt that reactivity is governed by the numbers of protons in nuclei. Vegetarians do not doubt that people shouldn't eat animals. Anti-abortionists do not doubt that the joining of a human sperm and a human egg creates a person with the same rights as any other person. In any such group, communication about issues of mutual concern proceeds on the basis of acceptance of certain basic postulates. Inquiry and discussion lead to new insights—factual insights in the case of groups sharing factual premises, ethical insights in the case of groups sharing ethical premises.

Ethical knowledge is also related to our emotional responses to events. Things that happen that are "not good" tend to disturb us. We condemn those who do these things and/or look for ways of setting things straight. Things that happen that are "good" tend to leave us undisturbed or glowing warmly. (The correspondence between our intellectual understandings and our emotional responses is, of course, neither precise nor perfectly controllable.) If we come to an understanding that something we thought was not good actually is, then there will be a tendency for us to be less disturbed by that event. Thus the search for ethical knowledge is in part an effort to regulate our emotional responses to events.

The Impact of Experience

Sometimes experience leads to an apparent conflict with theory. For example, shortly after Newton's theory of celestial mechanics was developed, it was noticed that although the theory explained the preponderance of planetary motion with great economy, it did not explain the precession of the perihelion of Mercury (the movement of the point of Mercury's closest approach to the sun). One might have said that this proved that the theory was wrong. But the theory was so successful in other ways that it was kept for three centuries before being classified as an approximation of relativistic theory that offered a better explanation of planetary motion.

The practice of keeping a theory even when it is demonstrably wrong, if there is no more attractive replacement, is evidence of our practicality and our need for something to believe in. It is understandable that a theory that is expected to give wrong answers in a very small percentage of cases should be seen to be more valuable than no theory at all, even if there is no way of predicting when the wrong answers will arise.

An old theory is replaced by a new one when the new one explains the phenomena that the old theory explained, resolves one or more difficulties encountered with the old theory, and appears likely to offer further interesting insights. The replacement of one theory by another is rarely instantaneous, unanimous, or free of pain, however. The new theory becomes established when a critical mass of people sees its virtues. Those who are first to see its possibilities seek fellow converts. But understandable conservatism keeps some from following. No theory can be rejected until it is first doubted, and doubt can be very discomforting. Theories in different areas of life—science, religion, politics, economics, ethics—have ties to one another, which make it hard to be confident that the replacement of a theory in one area will not necessitate replacements in other areas too. Every time a theory that had been asserted confidently is replaced, it is all the harder to assert any other theory confidently. And yet recrystallizing events occur, when suddenly we see the world differently, and we are at least as confident of the ultimate truth of our new views as we were of our old ones.

How Does Knowing Apply to Collective Decisions?

Knowing what a collectivity ought to do is much like other kinds of knowing. It can be as unshakable, even while it is limited to a particular group.

For collective decisions, as for the formation of individual intentions, knowledge is required that the outcomes that may be chosen are ethically permissible. Thus, for example, deciding to clear a forest would be precluded by knowing that the trees had a right to live. The sale of a daughter as a bride would

be precluded by knowing that each person should be free to choose his or her own marriage partner.

All collective decisions face the problem of inducing the members of the collectivity to form and maintain intentions to fulfill their designated roles in the coordinated plan, when:

a. Each might benefit personally by departing from the plan, and
b. Each might perceive that others would benefit by departing from the plan,

from which it might be concluded that it does not matter whether or not one fulfills one's assigned role, because others cannot be expected to fulfill their roles in any case.

To a substantial extent, the possibility of collective action is contingent upon the development of shared knowledge that the collective decision modes used by the collectivity identify good outcomes, and that therefore each member of the collectivity has an obligation to fulfill his designated role in the plan of coordinated actions.

There are four closely related but conceptually distinct arguments that may be used to keep the members of the collectivity from doubting that the selected outcome is good. Which one or more of the four are relevant will vary from one situation to another. The four arguments are:

1. The collective decision procedure provides evidence that the selected outcome is good.
2. The mode used is good, and therefore the selected outcome is good.
3. The selected outcome represents what has been agreed—either as outcome or as procedure, and it is good to keep agreements.
4. There is no possibility of achieving coordination by any means other than the procedure used. The selected outcome is good because it is better than the only real alternative of no coordination.

If an individual knows that the selected outcome is good, then the pull of conscience tends to offset the temptation of selfish interest. The knowledge that others are subject to the same pull of conscience helps create confidence that enough individuals will perform their planned roles to achieve the success of the effort at coordination, thereby inducing each individual to play his role even though his selfish interest would be better served by departing from the plan.

The following section examines the ways that the four arguments for the goodness of the outcome of a collective decision procedure apply to the eight modes of making collective decisions.

Applications to the Modes of Collective Decisions

Consensus

Consensus operates on a shared conviction that there is a correct outcome that the collectivity is seeking. They must know that the other members of the collectivity are not promoting their own interests, but are searching as well for the one proper outcome of their deliberations. Each member of the collectivity must know that it would be wrong for him to depart from the search for the right outcome and seek a selfish gain. If people give in to this temptation, the mode ceases to be consensus and becomes instead pseudo-consensus, trade or extortion.

Having defined the choice of an outcome as a problem of identifying the outcome that is objectively best, a collectivity operating on consensus can reasonably expect that if agreement on an outcome is achieved, there will be no problem of motivating individuals to form intentions to perform their designated roles. If all agree that the selected outcome is a correct one, then there should be no reason for any individual to depart from it. And if all understand that others agree with this, then there is no reason to expect them to decline to follow through and make one's own participation fruitless.

If there is a difficulty with achieving follow-through for consensus, it comes from the possibility that someone may see a reason for doing things differently after consensus has been reached. If the virtue of the selected outcome is that all agree to it, then its virtue no longer prevails when one person changes his mind. There may still, however, be some evidentiary value in the fact that everyone else was persuaded. The person who lost his conviction might persuade himself that if the discussion were reopened, the others would probably find some way to persuade him, and therefore suppress his doubts.

If doubts arise after agreement has once been reached, the idea that consensus is good will be telling only in conjunction with a belief that there is no alternative, that it is impossible, or at least too costly, to reopen the discussion of what to do. The goodness of consensus then lies in the opportunity of all in the collectivity to shape the outcome, rather than in their permanent belief in its correctness. In this context, the argument that agreements once made should be kept will have force as well.

Pseudo-Consensus

When a collective decision is reached through pseudo-consensus, evidence of the goodness of the selected option comes from the fact that the preponderance of the members of the collectivity, or at least those who feel strongly enough to twist the arms of the rest, are in agreement that the selected option is good.

Full consensus is generally regarded as superior to pseudo-consensus. However, if it is believed that full consensus is excessively costly or impossible to

attain, then pseudo-consensus may be regarded as the only available mode for making collective decisions, particularly if the number of persons involved is so large that trade is impractical. In these circumstances, pseudo-consensus tends to be regarded as not only the best available mode of making collective decisions, but also good in its own right.

Pseudo-consensus has some capacity to engender compliance from the idea that an outcome represents what was agreed, and therefore members of a collectivity should undertake their assigned roles in the intended collective action. However, the capacity of pseudo-consensus to do this is not as great as that of full consensus, because people who are prodded or cajoled into consenting to a collective decision have a noticeable propensity for deciding afterward that they will not do that which they agreed to do.

Trade

Trading partners accept that each seeks his own personal gain, so trade is more attainable than consensus. An outcome agreed by trade is not necessarily seen as correct, but only as agreed. However, the bargaining that precedes a trading agreement is a form of conflict, which can be detrimental to the trust that is needed between trading partners. To help preserve trust, trading partners often adopt the belief that certain elements of the agreement (particularly the price) are not subject to negotiation: The price is "the market price" or "a fair price." The idea of an external standard, which cannot be varied, can contribute to procedural efficiency as well as promote trust. No time need be spent arguing about things that "cannot be negotiated."

Trust is needed to avoid doubts about performance. For trade to occur, it is generally necessary for at least one of the trading partners (the one who performs second) to know that he is obliged to do what he agrees to do. This sense of obligation may come either from an ethical code or from an external agency (courts and police) that will enforce contracts. While it is not an absolute requirement, trade is greatly facilitated by a conviction on the part of each participant that the other participant(s) will carry out their obligations. To the extent that there is doubt about this, the incentive to reach trading agreements is weakened.

Trade is also dependent on shared knowledge concerning who owns what. If potential trading partners have different understandings of the validity of the ownership claims of the things traded, then the mode becomes extortion rather than trade.

The idea of the collective decision process generating evidence of the goodness of outcome has relatively little applicability to trade. Nor is the idea of there being no alternative particularly relevant: Renegotiation is possible more often than not. The ethical conception most powerful in forming and maintaining intentions to act in the way prescribed by a trade is that people ought to keep the agreements they

make. This is supplemented to some extent by a belief in the inherent goodness of trade as a way of coordinating intentions.

Extortion

For extortion, the conviction that is required is less than for trade, but achieving it is at least as difficult. Neither shared ethical knowledge nor agreement with respect to property rights is required. "All" that is required is that each first-performer have confidence that each second performer will behave as promised, and that each second-performer intend to do so. How this is achieved in the almost total absence of any incentive structure or ethical structure to promote performance is something of a puzzle.

The idea that people ought to keep agreements seems to have a slight impact for collective decisions made by extortion, perhaps a greater impact on those whose behavior is characterized as unethical than on those whose behavior is characterized as ethical: Kidnappers seem to release their victims when they promise more often than those who pay ransom keep promises not to involve the police. This is consistent with the general principle in law that agreements made under duress are not binding.

The idea that extortion is an inherently good mode, or that the process of reaching agreement identifies a good outcome has little applicability. There may be a slight applicability of the idea that once an extortionate agreement has been reached, no alternative coordination is possible, and the existing agreement is better than no coordination.

Authority

For authority, as for the agreement-on-procedure family generally, a principal focus of ethical support for collective decisions is the idea that the procedure used is good. This is often intertwined with the idea that there is no possibility of coordination in any other way. Belief in authority is often promoted by surrounding the authority with ceremony and by punishing behavior that would undermine confidence in authority. (Contempt of court is a punishable offense.) Confidence in authority is also promoted by making it "objectively" more credible, for example, by putting authority in the hands of persons who seem most likely to identify the best outcomes. The lack of an alternative way of achieving coordination arises naturally from the circumstances that lead to agreement on a procedure rather than agreement on an outcome. The perception that there is no alternative to the prevailing way of achieving coordination can lead people to find it attractive to mouth arguments that they do not believe, in favor of the status quo, on the ground that the status quo is better than the complete absence of coordination, which might result if too many people stopped believing in the status quo.

The practice of having judges write opinions in support of their decisions illustrates the idea that the outcome selected by authority can be shown to be good in the process of being selected. The opinion is supposed to convince any doubters that the chosen outcome is good.

For decisions made by authority, as for other modes in the agreement-on-procedure family, the idea that an outcome is good because it represents what was agreed may or may not apply. The possibility of the argument not applying comes from the fact that collectivities persist over time, and their membership changes. This still causes no problems if the collectivity is one that individuals join voluntarily. In joining, the individual agrees to what others have agreed to previously. However, if the collectivity is one that individuals are born into, then the idea of outcomes being good because they represent what was agreed is problematical. It is applicable only if there is an explicit ceremony of consent, once the person had reached an age of consent, or if being born is regarded as consenting to the institutions of the society into which one is born. In this case, the question of duress may arise. If a person has no "reasonable" alternative to remaining in the society into which he is born, then he can be said to have agreed to the rules of the society only in the sense in which a person agrees to turn over his wallet to a mugger. In practice, any concern about whether there has been consent to the rules of a society is diffused by conviction that the rules are good in any case.

Random Process

The outcomes chosen by random process are sometimes regarded as good by virtue of their having been chosen by super-human intervention. However, random process tends to be used when there is no generally agreed basis for asserting that one option is better than any other. In this case, what permits the outcomes selected by random process to be seen as good is the ability of random process to generate coordination when it might otherwise be impossible. More often than with other modes in the agreement-on-procedure family, the decision to use a random process is made on a case-by-case basis. In this event, the argument that the process is good because it was agreed pertains explicitly. The *ex ante* equality of the positions of individuals promotes agreement, and the idea that agreements ought to be kept maintains agreement once the outcome of the random process induces inequality in positions.

Voting

For collective decisions reached by voting, the idea that the process of determining the outcome generates evidence of the goodness of the outcome appears to have particular power. People often know that an outcome is good just because it is chosen by voting. This is intermingled with the idea of voting as objectively good.

There seems to be a strong tendency, at least in cultures of European origin, for people, upon meeting as a collectivity for the first time, to assume without discussion that collective decisions will be and ought to be made by voting. There seems to be a tendency to presume that voting is good because it reveals what ought to be done, that once a vote has been taken, the group knows what it ought to do.

The idea of there being no more attractive alternative is at work as well. There is great respect for consensus, but extreme patience and faith in the other members of a collectivity is required to employ consensus in all but the smallest groups. Trade fares no better in such groups: The possibility of multiple individuals holding out for disproportionate shares of the gains from trade makes it too likely that there will be no agreement. Extortion is generally not feasible because people have little in the way of effective threats against each other. Authority is feasible in a collectivity like a firm, where anyone who disagrees can leave and be replaced by someone willing to accept the decisions of the entrepreneur or chief executive officer. It is far less feasible in a collectivity that arises out of the possibility of mutually beneficial coordination by particular individuals: There is generally no natural way of determining who will exercise authority. Random process has some applicability, but voting would generally be favored for its perceived ability to identify the right outcome. Contest is difficult to employ because it requires prior agreement on the rules. Thus voting is in many circumstances understandably regarded as the best feasible mode for making collective decisions.

As with authority, the argument that voting is good because it was agreed to may or may not apply. It applies whenever it has been agreed that decisions will be made by voting. It does not apply to on-going collectivities that one joins without explicit agreement to do so, or under coercion.

Contest

Like consensus, authority and voting, contest can claim to identify the best outcome: The firm that ought to get the contract is the one that made the lowest bid. The student who deserves to be valedictorian is the one who achieved the highest grades. With the claim of identifying the best outcome goes the possibility of asserting that contest is objectively good. The idea that the outcome is good because it represents what was agreed to is usually applicable to contest as well. Our efforts to maintain discipline in this regard can be seen in the disrespect we have for “sore losers.” The idea of there being no alternative is generally not applicable to contest.

What Makes a Collective Decision Good?

The preceding section showed how all the modes of making collective decisions draw support within a collectivity from some combination of standard ideas about what is good. This section seeks to identify external criteria for goodness in collective decisions.

In the preceding section, something could be said to be good if a critical mass of the members of a collectivity regarded it as good. In this section, the reference group is not a collectivity but rather the readers of this book. Thus for this section, the test of whether something is good is whether a critical mass of my readers regard it as good.

I propose that a collective decision is good if the following three conditions are met:

1. The decision is made by a procedure that is accepted by the members of the collectivity.
2. The outcome does not conflict with the reasonable claims of persons who are not members of the collectivity.
3. Any member of the collectivity who disagrees with the outcome can leave on reasonable terms.

These are not conditions whose satisfaction might be established indisputably. Rather, they are an agenda for discussion of whether a collective decision is good. Thus, what I am proposing is that inquiry into whether collective decisions are good should address these questions. If a group is persuaded that the questions can be answered affirmatively, then they have reason to call the decision good and to anticipate positive or neutral emotional responses to the decisions. If one or more of the questions can only be answered negatively, then the basis for calling the decision good is missing. Negative emotional responses can be anticipated, and those who come to such conclusions might ask themselves whether there is some ameliorative action that they ought to take.

The significance of the conditions is discussed in the following subsections.

Acceptance of Procedures by the Collectivities that Use Them

This condition is satisfied almost automatically. If a collectivity has made a decision, it has almost certainly accepted the procedure by which the decision was made. One significant aspect of this condition is that it emphasizes the fact that what must be achieved is acceptance by the members of the collectivity, and not by outsiders. If someone who is not a member of the collectivity is dissatisfied with the procedure, then the relevant question is, does the decision fail to respect the reasonable claims of those who are not members of the collectivity? This is taken up in the next subsection.

If some members of a collectivity are not satisfied, then one might ask first whether there are reasonable opportunities for them to leave the collectivity if they wish. This issue is addressed in the third subsection. If they do not wish to leave the collectivity but are dissatisfied with a procedure, then the question of whether the procedure is a good one is a political question for the collectivity. Only time can tell whether those who are dissatisfied will prevail.

Respect for the Reasonable Claims of those Outside of Collectivities

This condition is illustrated by the idea that if a collectivity decides to terrorize and plunder some individual or group outside the collectivity, the fact that the decision was reached by consensus within the collectivity that plans to terrorize and plunder does not make the decision good. People have claims that deserve the respect of others. Defer for the moment the question of what those claims are. Some would call such claims "rights." However, the word "rights" conveys a connotation of an enforcement structure that will do whatever is necessary to insure that the rights are not violated. To separate the question of any such deployment of power, and to emphasize that what is sought is agreement among those who wish to discuss the subject, let the word "claims" be used.

Not everyone agrees that claims exist outside of collectivities. Hobbes regarded the absence of agreement as a state of "warre of every man against every man."

> To this warre of every man against every man, this also is consequent; that nothing can be Unjust. The notions of Right and Wrong, Justice and Injustice have there no place. Where there is no common Power, there is no Law: where no Law, no Injustice. Force, and Fraud, are in warre the two Cardinall virtues … It is consequent also to the same condition that there be no Propriety, no Dominion, no *Mine* and *Thine* distinct; but onely that to be every mans, that he can get; and for so long, as he can keep it.[4]

The Hobbsian view that claims that ought to be respected (or rights) do not exist in the absence of agreement is endorsed in this era by James Buchanan:

> As Thomas Hobbes perceptively noted, in this state of nature each person has a "right" to everything … [A]s a result of the actual or potential conflict, … some "natural distribution" will come to be established. This cannot be classified as a structure of *rights*, since no formal agreement is made, …[5]

This view denies the innate morality of people. In fact, people recognize obligations to respect the reasonable claims of others, apart from agreements to do so. Thus the state of nature is closer to that proposed by Locke:

> … we must consider what State all Men are naturally in, and that is, a *State of perfect Freedom* to order their Actions, and dispose of their Possessions, and Persons as they

> think fit, within the bounds of the Law of Nature, without asking leave, or depending upon the Will of any other Man.
>
> A *State* also *of Equality*, wherein all the Power and Jurisdiction is reciprocal, no one having more than another …
>
> …
>
> The *State of Nature* has a Law of Nature to govern it, which obliges every one: And Reason, which is that Law, teaches all Mankind, who will but consult it, that being all equal and independent, no one ought to harm another in his Life, Health, Liberty, or Possessions.[6]

The corresponding requirement for relations among independent collectivities is that no collectivity encroach upon the territory of another or interfere in other ways. But this requires a shared understanding of what constitutes "the territory of a collectivity" and what constitutes interference. In practice, there are bound to be disagreements. If one collectivity wishes to build a nuclear power plant, persons outside the collectivity may feel that this exposes them to an unwarranted risk. Facts will be disputed. The best that we can hope for is that we will be able to muddle through these difficulties with good faith.

Reaching a consensus regarding what constitutes "the territory of a collectivity" is far from a simple matter. It is not reasonable to expect a consensus for freezing the status quo of any particular time, when that status quo has been shaped extensively by violence.[7] A collectivity's claim to territory should be based not just on history, but also on evidence that the fraction of the natural endowment that is claimed is no greater than the fraction of the world's population that makes the claim. The assertion of control over a disproportionate share is an infringement on the reasonable claims of others. Therefore a collective decision that entails such a disproportionate claim should not be described as good.

Provision of a Reasonable Exit Option

An obvious aspect of this condition is illustrated by the injustice we perceive when a country does not permit its citizens to emigrate. The general condition is that collectivities should be composed of individuals who coordinate their actions and their intentions *voluntarily*.

Within the idea of collectivities formed by voluntary association, there is a question of the extent to which individuals should be able to commit themselves to collectivities. Suppose there were a club, formed by voluntary association, with a rule that anyone who sought to leave the club would be killed. Suppose there were a country with a law that their citizens would be asked to take oaths of allegiance when they attained the age of 18. Those who did not wish to do so would be required to depart, and those who took the oath would not be permitted to depart thereafter. Would it be good to implement such a rule in a club or such a law in a nation? No. While there is value in the keeping of commitments, the ways in which people are penalized for breaking commitments should be limited to the

termination of reciprocal commitments and the loss of honor and accumulated wealth. Each person should always have the option (though there might be reasons of honor for not doing so) of severing his connections with any and all collectivities of which he is a member and claiming his share of the future flow of what nature provides as a member of the collectivity or collectivities that he now chooses, or as a hermit.

Summary

This chapter began with a discussion of what knowing is. The traditional definition of knowing is that a person knows something if: 1. it is true; 2. the person believes that it is true; and 3. the person has a good reason for his belief. It was argued that the procedures that we use to seek to determine whether things are true are fallible, and yet we use "know" in an understandable way. Therefore it must be that it is not necessary for something to be true in order for it to be known. It is necessary only that the person in question believe it is true, and that he have a reason for his belief that is accepted by those with whom he communicates.

Ethical knowledge, it was argued, has the same structure as factual knowledge. A person knows that something is good if he believes that it is good and has a reason for his belief that is accepted by the persons with whom he communicates.

Ethical knowledge shapes collective decisions first in limiting the agenda of permissible outcomes. It also helps to provide the cohesion needed to induce individuals to form and maintain intentions to perform their assigned roles in the chosen outcomes. Four different forms of ethical argument for the goodness of collective decisions were noted:

1. The collective decision procedure produces evidence that the selected outcome is good.
2. The mode used is good, and therefore the selected outcome is good.
3. The selected outcome represents what has been agreed—either as outcome or as procedure, and it is good to keep agreements.
4. There is no possibility of achieving coordination by any means other than the procedure used. The selected outcome is good because it is better than the only real alternative of no coordination.

These arguments are used in different combinations in support of collective decisions that are reached by the eight identified modes.

For the evaluation of collective decisions by persons who are not members of the collectivity making the decisions, a different conception of good is needed: The good is that which is agreed to be good by those who are making the evaluation. For this perspective, I propose that the following questions be asked:

1. Is the decision made by a procedure that is regarded as good by the members of the collectivity?

2. Does the outcome conflict with the reasonable claims of those who are not members of the collectivity?

3. Does every member of the collectivity have the opportunity to leave the collectivity on reasonable terms if he wishes?

If these conditions are satisfied, decisions deserve our approval. If not, we may wish to ask whether there is some action we ought to take to promote their satisfaction.

There is an important difference between this ethical framework and the ethical frameworks that emerge from the writings of Pareto, Bergson, Samuelson, Hicks and Kaldor discussed in Chapter 3. All of these employ an implicit assumption that the collectivity under consideration has an immutable membership. People are tied in compulsory association by some unstated principle. This presumption of compulsory association leads to a need to justify whatever rules or decisions are imposed on the collectivity by less than unanimous agreement. This tends to lead to absolutism in implicit epistemologies: To justify imposing anything on people who wish you wouldn't, one must have a structure that guarantees that the imposed decisions are objectively correct.

In contrast, the epistemology advanced in this book is explicitly "relative." Knowing is relative to a communicating group. Compulsory association is reduced to the bare minimum that arises unavoidably from sharing existence on a single planet. The necessity of muddling through without a guarantee of being right is recognized for the interdependencies characterized by compulsory association.

The ethic advanced by Buchanan and Tullock in *The Calculus of Consent*, with its emphasis on agreement, has much in common with that advanced in this book. However, the ethic advanced here draws attention to the needs for respect for the reasonable claims of those who are not members of the collectivity and for an explicit exit option, in addition to the need for agreement within a collectivity.

Reasons for an In-Depth Analysis of Voting

This ends the general discussion of collective decisions. The remainder of the book concentrates on collective decisions that are made by the mode of voting.

Voting is interesting not because it yields outcomes that can be demonstrated to be good or right, but rather because collectivities often find it to be the most attractive available mode for deciding what they shall do. Consensus is more attractive when it can be attained at a reasonable cost, but in large groups it usually cannot. The other modes in the agreement-on-outcome family suffer from the same difficulty, though to a lesser extent. Pseudo-consensus and trade can sometimes be achieved when consensus cannot, but there can be no guarantee that these will be attainable. Making agreements by extortion is virtually never mutually attractive to the members of a collectivity. If the members of a collectivity agree that they wish to coordinate their intentions in some way and

only need to settle on which way, then they look to modes in the agreement-on-procedure family.

The mode of authority can be attractive when unusual competence is readily identified and it can be expected that authority will not be abused. But authority often is abused, and even when it is not, the glaring lack of symbolic equality that is entailed in authority makes it often unacceptable. Random process can be attractive for its symbolic equality and high procedural efficiency when there is no agreed basis for finding one option more desirable than others, but its severe deficiency in outcome efficiency makes it generally unattractive. Contest can be attractive when it is possible to reach agreement on how the contest ought to be judged. However, the members of a collectivity often disagree about this. Voting is the one mode that is left.

Voting is extremely attractive for its symbolic equality and procedural efficiency. The most serious weakness of voting is with respect to the criterion of preservation. Collectivities overcome this weakness through the use of constitutions that limit the range of issues that can be decided through voting.

To get as much as possible from voting, it is important that the strengths and weaknesses of its numerous variations be understood. Developing these is the task of Part II of the book.

Notes

1 For a discussion of knowing in these terms, see Robert Nozick, *Philosophical Explanations* (Oxford, 1981), Chapter 3.
2 For an exposition of Bayes' Theorem, see B. W. Lindgren, *Statistical Theory* (New York, 1962), pp. 22–6.
3 Tad Tuleja, *Fabulous Fallacies* (New York, 1982), p. 1.
4 Thomas Hobbes, *Leviathan, Parts I and II* (Indianapolis, 1958), Chapter 13, p. 108.
5 James Buchanan, *The Limits of Liberty: Between Anarchy and Leviathan* (Chicago, 1975), p. 24.
6 John Locke, *Two Treatises of Government* (New York, 1960), Second Treatise, sections 4 and 6, pp. 309, 311. Robert Nozick, in *Anarchy, State, and Utopia* (New York, 1974), elaborates a framework based on Locke's state of nature.
7 For a fuller development of these ideas see Nicolaus Tideman, "Commons and Commonwealths: A New Framework for the Justification of Territorial Claims," in Robert Andelson (ed.) *Commons Without Tragedy* (London, 1991). Bruce Ackerman, in *Social Justice and the Liberal State* (New Haven, 1980), presents a framework with many similarities.

PART 2
Voting

This page intentionally left blank

Chapter 7

An Overview of Voting

Introduction

Expanding on the definition in Chapter 2, a procedure for making collective decisions by voting (a "voting procedure") is composed of the following five elements:

Setting the Agenda

Some agreed process is used to identify a set of options for consideration. When voting is used to fill an office, the options are the nominated candidates. In other circumstances the options are alternative courses of action. The word "options" will be used to denote the possibilities among which a voting process selects, whether the options are candidates or more general courses of action.

Eliciting Evaluations

Each member of the collectivity provides evaluative information with respect to the options, in some prescribed form. In the type of voting that is most commonly used, each voter specifies the single option for which he or she votes. The information required of each voter could be a ranking of the options or a set of scores for the options, or it could take yet other forms. The evaluative information, whatever form it may take, will be referred to hereafter as a "vote."

Processing the Evaluations

The votes are processed in a way that can be described by a mathematical function. A function used for this purpose will be called a "vote-processing rule." The vote-processing rule yields scores for the options or a ranking of the options or in some other way identifies one of the options (or more if the result is "a tie") as the winner(s) of the election.

Dealing with Ties

If the result is a tie, either there is no collective decision or else an additional procedure (generally involving authority or random process) is used to break the tie.

Inducing Intentions

The result is announced, and this creates in members of the collectivity intentions corresponding to the result. Thus a collective decision has been made.

The remainder of the book deals with the possible attractiveness of different forms of steps 2, 3, and 4 above, in various situations in which voting might be used. The next section classifies these situations.

The Dimensions of Voting Situations

Voting situations differ in four principal dimensions: the presumed interestedness of the voters, the number and spatial configuration of the options, the weights to be attached to the votes, and the number of options to be selected. The remainder of this section discusses these dimensions one by one.

The Presumed Interestedness of Voters

Voters may be understood to be either disinterested judges of what is best or advocates of their self-interests. The voting by judges in beauty pageants or in aesthetically oriented athletic events such as diving or gymnastics illustrate voting by disinterested judges. Democracy is sometimes idealized as a process in which votes express the judgments of voters as to what promotes the good of the whole, although it is probably more common to place a more selfish interpretation on democratic votes. Voters are often referred to as "voting their pocketbooks."

When voters are viewed as self-interested advocates, it is important to consider the possibility that the votes may reflect not the true views of voters, but rather their guesses as to what permitted votes will best advance their interests. For example, if three options are under consideration in an election in which each voter may cast a vote for one candidate, a voter may vote for his second choice rather than his first if he believes that his first choice has no chance of winning. On the other hand, if voters are understood to be acting as disinterested judges, then it is presumed that their votes represent honest expressions of their opinions rather than strategic calculations.

Number and Configuration of Options

The simplest case with respect to this dimension of voting is a choice between just two options. When a choice is to be made among three or more options, there are significant complications that do not arise under a choice between two options. A third possibility of interest occurs when the options under consideration form a continuum of a single dimension. The choice of a speed limit exemplifies this possibility. A fourth possibility is that the options form a continuum of two or more dimensions. A decision about how much to budget for each of a variety of programs exemplifies this possibility. A fifth and final possibility is that the options, while discrete, can be located in a continuum of one or more dimensions. This possibility may be exemplified by a choice of a computer for a university, when there is a specific number of options, each of which can be described by its price and speed, and the options do not differ in any other way.

Weights for Votes

There are three possibilities of interest with respect to weights for votes. The first and most common case is that all votes are weighted equally. The second case, exemplified by votes among the stockholders in a corporation, involves assigning to each vote a weight that the voter cannot affect by the vote that he or she casts. The case will be referred to as the case of predetermined unequal weights. The third case combines elements of the first two. It is like the first case, in that all voters have the same voting opportunities, but like the second, in that voters can affect the weights given to their votes. This case may be exemplified by the following process: A collective decision is to be made between two options. Each voter assigns to the option he or she favors a score representing the voter's statement of how much more valuable the favored option is than the alternative. The overall score of each option is the sum of the scores assigned to it, and the option with the greater score is chosen. While this particular procedure is not very attractive, there are other procedures in this category (sometimes called "demand-revealing processes") that may be. Procedures in this category will be referred to as having endogenously determined unequal weights.

Number of Options to be Selected

There are two cases of interest in this dimension. The first and most common case is the selection of one option. The second case includes the selection of any number of options greater than one. Systems of proportional representation exemplify this case. It is possible to treat the selection of more than one option as the selection of a single "revised option," where each "revised option" is a set of original options. However, because of the possibility of using votes with respect to

single options as the basis for selection among revised options, it is appropriate to treat this case separately.

Summary of Dimensions

The dimensions in which voting situations differ are summarized in Table 7.1.

Table 7.1 The dimensions of voting situations

Dimension	Possibilities
Interestedness	Disinterested voters
	Self-interested voters
Number and configuration of options	Two options
	Three or more discrete options
	A continuum of options in a single dimension
	A continuum of options in two or more dimensions
	Discrete options located in a continuum
Weights on votes	Equal weights
	Unequal predetermined weights
	Unequal endogenously determined weights
Number of options to be selected	One
	Two or more

Table 7.1 serves several practical functions. First, if one is seeking a voting procedure for particular circumstances, locating the circumstances in terms of the dimensions can help clarify which procedures may be attractive. Second, if one has an idea for a new voting procedure, considering the dimensions may clarify the circumstances in which the procedure will be attractive. Finally, if one is interested in making generalizations about voting procedures, thinking about the dimensions may clarify the circumstances in which the generalization is valid.

The remaining chapters of the book can be classified in terms of the combinations of dimensions with which they deal. Chapter 8 is concerned with rationales for using majority rule to choose one out of two options, for voters who may be either disinterested judges or self-interested advocates. The analysis is concerned primarily with equally weighted votes, but votes with unequal predetermined weights are also considered.

Chapters 9 through 13 are concerned with situations in which there are more than two options under consideration. The number of options to be selected remains one, the weights may be either equal or unequal and predetermined, and the voters may be either disinterested judges or self-interested advocates. Chapter 9 is concerned with the significance of the possible appearance of majority-rule cycles. Chapter 10 provides a exegesis of the Arrow Theorem, which shows that,

because of the possibility of cycles, it is not possible to guarantee consistency in a reasonable way when the set of options under consideration expands or contracts. Chapter 11 is concerned with the Gibbard–Satterthwaite theorem, which pertains only to self-interested voters and shows that when there are more than two options, such voters cannot generally be motivated to report accurately their rankings of the options under consideration. While the Arrow theorem and the Gibbard–Satterthwaite theorem specify properties that voting procedures cannot attain, Chapter 12 develops criteria that voting procedures can attain, criteria by which vote processing rules can be evaluated. Chapter 13 evaluates alternative procedures in terms of the properties developed in Chapter 12.

Chapter 14 changes the focus to situations where the options lie in continuums. The cases of a one-dimensional continuum, a continuum of two or more dimensions, and discrete options in a continuum are each treated. The number of options to be selected remains one; voters may be either disinterested judges or self-interested advocates, and votes may have either equal or unequal predetermined weights. Having the options in continuums makes new voting procedures possible.

Chapter 15 relaxes the assumption that only one option is to be selected. The most interesting new possibilities here pertain to self-interested voters who are to be represented "proportionally" in a deliberative body. Generally, the options cannot be represented on a continuum. Votes may have either equal or unequal predetermined weights.

Chapter 16 moves into the realm of votes with endogenous unequal weights, when voters are self-interested advocates. It is assumed in this Chapter that only one option is to be selected, and all possibilities with respect to the number and configuration of options are considered.

This page intentionally left blank

Chapter 8

Majority Rule and its Weighted Analog

Introduction

"Majority" means "more than half." Majority rule is well defined only for cases where there are just two options. For these cases, majority rule is the vote-processing rule that chooses the option that has the greater number of supporters (the majority of those voting). If there are more than two options, it is possible that no option has the support of a majority (more than half) of the voters. Vote-processing rules for these cases are considered in later chapters.

In terms of the dimensions of voting developed in Chapter 7, majority rule is characterized by the fact that it involves two options, with one to be selected. Traditionally, it also involves equally weighted votes. This standard version of majority rule is treated in the next section, where separate subsections consider majority rule historically, as a procedure for estimating statistically which option is likely to be better, and as an axiomatic system. The statistical subsection accords separate treatment to majority rule for voters who are self-interested advocates and for voters who are disinterested judges. The next section discusses the variation on majority rule generated by the assignment of predetermined, unequal weights to the votes. Here again, separate treatment is accorded to the use of this rule when voters are disinterested judges and when they are self-interested advocates.

Equally Weighted Votes

Majority rule is so much a part of our cultural heritage that it is difficult to address the question of what is attractive about it. Should we have minority rule instead? Absurd.

Historical Approaches

For a systematic explanation of why majority rule is attractive, it is interesting to examine theories of its origin. J.A.O. Larsen argues that a comparison of decision-making procedures described in the Homeric epics with what is known about the origins of democracy in Greece suggests that voting originated in the seventh century B.C.[1] He says that the earlier practice was simply to talk until it seemed

that someone had come up with the right idea. But this has a potential for leading to conflict when not all persons perceive the discussion to have come to the same conclusion. Larsen offers such an example from the second book of the Iliad. Agamemnon planned to have others oppose his insincere suggestion that they all go home. The men responded by rushing to the boats, and Odysseus had to drive them back to the meeting.[2] An orderly discussion followed by a vote can help avoid such disorder.

Another theory holds that voting originated independently in Germanic tribes of Northern Europe, in an environment where it was understood that there should be unanimous agreement on all decisions.[3] If consensus did not emerge through discussion, it was acceptable to employ coercion to reach a decision to which all agreed, with those who did not wish to be coerced allowed to flee. Majority rule emerged when it became customary to allow the majority (who could be expected to prevail in the event of a clash) to have their way without a physical clash.

Statistical Justifications

Self-Interested Advocates There is a statistical justification of majority rule in terms of the "principle of insufficient reason." If one is trying to estimate which of two options is likely to provide greater total utility to the members of a collectivity, and if the only information that is available is the number of voters who favor each option, then there is no reason to suppose that those who favor one option will feel any more strongly about the matter than those who favor the other. Therefore it is more likely that greater total utility is obtained if the option favored by the majority is adopted than if the option favored by the minority is adopted.

Disinterested Judges If one views voters as disinterested, then a justification of majority rule first offered by Condorcet is possible.[4] If every voter is regarded as equally skillful in discerning which of two options is better, and if prior to the vote one has no reason for believing that one option is more likely to be better than the other, then after the vote it is reasonable to believe that the option favored by the majority is more likely to be the better one.

May's Axiomatic Justification

A different way of justifying majority rule is to look for a set of formal, mathematical conditions that describe an attractive vote-processing rule and show that they are uniquely satisfied by majority rule. This approach was taken by K.O. May.[5] May characterized the object of a vote as a "motion" and used the numbers "1," "0" and "–1" to symbolize a vote in favor, an abstention, and a vote against the motion respectively. In describing the outcome of the vote, he used the same numbers to symbolize "the motion is passed," "there is a tie," and "the motion is defeated" respectively. In May's framework, a "voting rule" (what I have called a

vote-processing rule) is a mathematical function that maps a sequence of numbers, each of which is 1, 0, or –1, to one of the numbers 1, 0, and –1. Thus if S is a sequence of 1's, 0's and –1's, majority rule may be described as the following function:

$$f(S) = \begin{cases} 1 \text{ if the sum of the components of } S \text{ is positive;} \\ 0 \text{ if the sum of the components of } S \text{ is 0;} \\ -1 \text{ if the sum of the components of } S \text{ is negative.} \end{cases}$$

May showed that the following four conditions are necessary and sufficient for majority rule:

1. *Universal domain*: The function yields exactly one outcome for every sequence of 1's, 0's and –1's interpreted as votes.
2. *Anonymity*: The outcome does not depend on who cast which vote. Formally, if f is the function that describes the vote-processing rule, S is a sequence of 1's, 0's and –1's, and S' is a permutation of S, then $f(S') = f(S)$.
3. *Neutrality*: If everyone changes his vote to the negative of what it had been, then the outcome changes to the negative of what it had been. Formally, if S is a sequence of 1's, 0's and –1's, and S' is a sequence of the same length in which each component is the negative of the corresponding component of S, then $f(S') = -f(S)$.
4. *Positive responsiveness*: Other things being equal, when a voter's opinion of a winning motion improves, the motion remains a winning motion, and when a voter's opinion of a tied motion improves, the motion becomes a winning motion. Formally, if $f(S) \geq 0$ and every component of S' equals the corresponding component of S except for one which is greater than the corresponding component of S, then $f(S') = 1$.

It may seem strange that positive responsiveness is defined only in terms of improvements in a voter's opinion of winning and tied motions when it is equally important that a losing motion not become a tie or winning motion by the addition of one more vote against it. This is not an oversight, however. If a vote-processing rule is neutral, then any violation of this "negative" version of positive responsiveness implies a violation of positive responsiveness as defined.

May's conditions are *necessary* for majority rule because, as can easily be seen, they are all satisfied by majority rule. The proof that they are *sufficient* for majority rule is as follows: Suppose that a vote-processing rule satisfies the four conditions. Because the rule is anonymous, it depends only on the number of votes that are 1, 0, and –1, and not on where in a sequence of votes these numbers appear. Define $N(1)$ to be the number of 1's, $N(0)$ to be the number of 0's, and $N(-1)$ to be the number of –1's in a sequence, S. Consider a case where $N(1) =$

$N(-1)$ in S. The assumption that the vote-processing rule satisfies the condition of universal domain implies that $f(S)$ is defined. It cannot be that $f(S)$ is 1, because if it were, then when all the votes were changed to their opposites $f(S)$ would be –1 by neutrality. But this cannot be, because $f(S)$ depends only on $N(1)$, $N(0)$ and $N(-1)$, and these have not changed, so $f(S)$ cannot have changed from 1 to –1. For the same reason, $f(S)$ cannot be –1, so it must be 0 because this is the only other possibility. Thus $f(S) = 0$ when $N(1) = N(-1)$.

Now suppose $N(1) = N(-1) + 1$. If one of the votes that is 1 is changed to 0, then $f(S)$ is 0 by the previous argument. Moving from that situation to where $N(1) = N(-1) + 1$ is an example of the premise of positive responsiveness. Therefore the conclusion of positive responsiveness applies, and $f(S) = 1$ when $N(1) = N(-1) + 1$. In the same way, one can move from $N(1) = N(-1) + 1$ to $N(1) = N(-1) + 2$, to $N(1) = N(-1) + 3$, and so on, so that $f(S) = 1$ whenever $N(1) > N(-1)$. From neutrality it follows that $f(S) = -1$ whenever $N(1) < N(-1)$. The three conclusions reached,

$$f(S) = 0 \text{ whenever } N(1) = N(-1)$$
$$f(S) = 1 \text{ whenever } N(1) > N(-1)$$
$$f(S) = -1 \text{ whenever } N(1) < N(-1)$$

describe majority rule. Thus the conditions of universal domain, anonymity, neutrality and positive responsiveness are sufficient to ensure that a vote-processing rule for pairs of options is majority rule.

May also showed that the conditions are independent, by giving an example for each condition of a rule that fails to satisfy that condition while satisfying the other three. For universal domain, May's example is a rule in which the outcome is both 1 and –1 when $N(1) = N(-1)$. For anonymity, the example is a rule in which the first voter's vote counts double. For neutrality, it is a rule in which a vote of more than two thirds is required to pass a motion. For positive responsiveness, it is the rule that there is a tie unless the vote is unanimous.

May's theorem permits one to defend majority rule in the following way: "I would not want a vote-processing rule for pairs of options that did not possess the properties of universal domain, anonymity, neutrality, and positive responsiveness. Since majority rule is the only vote-processing rule for pairs of options that has these properties, majority rule is the only vote-processing rule for pairs of options that I could want."

A Modification of May's Theorem[6]

The condition of positive responsiveness is reasonable enough when there are just two options, but its natural extension to more than two options is less appealing. The extension of positive responsiveness to multiple options would be that if there is a tie, and any voter raises his ranking of a tied option in any way, even from last to second-to-last, this results in that option being the sole winner. Because it is

somewhat doubtful that people would want to commit themselves to this condition, and because the set of vote-processing rules for multiple options that satisfy this condition is quite limited, it is interesting to characterize majority rule in a different way. This alternative characterization requires the introduction of the conditions of monotonicity[7] and resolvability.

Monotonicity in a vote-processing rule for pairs of options is the condition that, other things being equal, when a voter improves his opinion of a winning motion, it remains a winning motion, and when a voter improves his opinion of a tied motion, either it remains a tied winner or it wins. Formally, if every component of $f(S')$ equals the corresponding component of $f(S)$ except for one which is greater than the corresponding component of $f(S)$, then $f(S') \geq f(S)$.

Resolvability in a vote-processing rule for pairs of options is the condition that, if there is a tie, then, for each option, x, in favor of which the tie could be resolved, there is a vote, v, that resolves the tie in favor of x when v replaces an abstention. More formally, if $f(S) = 0$, $| x | = 1$, and S_i (component i of S) is 0, then there is some S' such that every component of S' except for S_i is the same as the corresponding component of S, and $f(S') = x$.

The modification of May's theorem replaces the condition of positive responsiveness with the two conditions of monotonicity and resolvability.

It is clear that majority rule satisfies the conditions of monotonicity and resolvability. The basic pattern of May's proof is followed by the proof that the five conditions of universal domain, anonymity, neutrality, monotonicity and resolvability are sufficient to determine that a vote-processing rule for pairs of options is majority rule. Because the rule is anonymous, it depends only on $N(1)$, $N(0)$ and $N(-1)$. Because it possesses universal domain and is neutral, $f(S) = 0$ when $N(1) = N(-1)$.

Now suppose $N(1) = N(-1)$ and $S_i = 0$. By resolvability, there is some change in S_i that yields an outcome of 1. Suppose that changing S_i to -1 yielded an outcome of 1. Then changing S_i from -1 back to 0, thereby changing the outcome from 1 to 0, would be a violation of monotonicity. Therefore, since it is the only other possibility, it must be that changing S_i to 1 yields an outcome of 1. Thus $f(S) = 1$ when $N(1) = N(-1) + 1$. In the same way, one can move from $N(1) = N(-1) + 1$ to $N(1) = N(-1) + 2$, to $N(1) = N(-1) + 3$, and so on, so that $f(S) = 1$ whenever $N(1) > N(-1)$. From neutrality it follows that $f(S) = -1$ whenever $N(1) < N(-1)$. Thus majority rule is again described. Thus the conditions of universal domain, anonymity, neutrality, monotonicity and resolvability are sufficient to describe majority rule.

The examples that May gave to show that universal domain, anonymity and neutrality are independent conditions continue to apply to the modified theorem. That resolvability is an independent condition is proved by the fact that the unanimity rule possesses the other four conditions of the modified theorem but not resolvability. That monotonicity is an independent condition is proved by the fact

that the opposite of majority rule (the option with the fewer votes wins) possesses the other four conditions of the modified theorem, but not monotonicity.

The modification of May's theorem shows that another way of defending majority rule is by saying that majority rule is the only vote-processing rule for pairs of options that satisfies the conditions of universal domain, anonymity, neutrality, monotonicity and resolvability, and these are conditions that would be possessed by any voting rule that one would want.

Predetermined Unequal Weights

Voting is generally regarded as an institution characterized by "one person one vote." And yet it is possible to employ voting in procedures where unequal weights are applied to the votes cast by different voters, as is done, for example, in stockholders' meetings. There is one rationale for weighting the votes when the voters are viewed as disinterested judges, and a different rationale when voters are viewed as self-interested advocates.

Disinterested Judges

When voters are viewed as disinterested judges, their role is to select from the available options the one that best satisfies a criterion of good that is shared by a collectivity, and yet not directly measurable in the way that weight, for example, is measured. There may be aesthetic judgments involved, or the measurement difficulty may be derived from future consequences that cannot be observed when the choice is made. Whatever the source of the measurement problem, there may be reasons for believing that some voters are better judges than others. In the case of aesthetic judgments, some judges may have a record of identifying the choice that is regarded as best by a consensus of other judges. In the case of uncertain future consequences, some judges may have a record of predicting these consequences accurately. In either case, the differing abilities of judges would provide a reason for believing that weighting the votes unequally, with more weight going to the better judges, could be expected to generate better outcomes on average than equal weights for all voters. An application of Bayes' theorem, which was mentioned in Chapter 6, shows the way in which the weights applied to different judges should depend on their ability.

Recall from Chapter 6 that Bayes' theorem states that if there are two possible states of the world, A (option A is better) and B (option B is better), and if, from some unspecified source, the probabilities of the two states prior to the incorporation of the implications of data D are P_A and P_B respectively, and if the probabilities of observing data D given the two states are $P_D \mid A$ and $P_D \mid B$ respectively, then the implication of observing data D is that the probabilities of A and B become

$$\frac{P_A \bullet P_D \mid A}{P_A \bullet P_D \mid A + P_B \bullet P_D \mid B} \quad \text{and} \quad \frac{P_B \bullet P_D \mid B}{P_A \bullet P_D \mid A + P_B \bullet P_D \mid B} \tag{8.1}$$

respectively.

Suppose there are N voters, of whom k (numbered 1 through k) believe that state A is true, while the remainder believe that state B is true. And suppose that for each voter, i, there is a known probability from previous experience, P_i, that that voter will discern the correct state. Suppose further that the probabilities that different voters will discern the correct state are independent. Then the probability of the observed votes if the state of the world is A is

$$P_D \mid A = \left[\prod_{i=1}^{k} P_i\right]\left[\prod_{i=k+1}^{N}(1-P_i)\right]. \tag{8.2}$$

(The expression $\prod_{i=1}^{k} P_i$ is mathematical notation for $P_1 \bullet P_2 \bullet P_3 \bullet \ldots \bullet P_k$. More generally, the symbol $\prod_{i=j}^{k}$ denotes the product of cases of the expression that follows, from the case where $i = j$ and then increasing i by 1 for each case until $i = k$.) The probability of the observed votes if the state of the world is B is

$$P_D \mid B = \left[\prod_{i=1}^{k}(1-P_i)\right]\left[\prod_{i=k+1}^{N} P_i\right]. \tag{8.3}$$

Substituting these expressions into equation 8.1 permits calculation of probabilities of the two states of the world, given the votes. However, it is not necessary to compute the probabilities to know which state is more likely. The expressions for the probabilities of the two states have the same denominators, so these can be ignored in determining which probability is greater. Furthermore, the assumption that prior to the vote neither state can be regarded as more likely than the other is the assumption that $P_A = P_B = 0.5$, so these terms have no bearing on which state is more likely. Thus the more likely state is determined entirely by whether $P_D \mid A$ or $P_D \mid B$ is greater. Thus A is the more likely state if and only if

$$\left[\prod_{i=1}^{k} P_i\right]\left[\prod_{i=k+1}^{N}(1-P_i)\right] > \left[\prod_{i=1}^{k}(1-P_i)\right]\left[\prod_{i=k+1}^{N} P_i\right], \tag{8.4}$$

which occurs if and only if

$$\prod_{i=1}^{k} \frac{P_i}{1-P_i} > \prod_{i=k+1}^{N} \frac{P_i}{1-P_i} . \tag{8.5}$$

This inequality is preserved when the natural logarithms of both sides are taken, so that *A* is the more likely state if and only if

$$\sum_{i=1}^{k} \ln\left(\frac{P_i}{1-P_i}\right) > \sum_{i=1}^{k} \ln\left(\frac{P_i}{1-P_i}\right) \tag{8.6}$$

(The mathematical symbol Σ has a meaning parallel to that of Π, but with the terms summed rather than multiplied together.) The ratio $\frac{P_i}{1-P_i}$ is the odds that voter *i* will discern the state of the world correctly. Equation 8.6 says that to determine which state is more likely (that is, to determine which option is more likely to be better), give each state a score equal to the sum of the natural logarithms of the odds that the believers in that state are right. The state with the higher weighted sum is the more likely one. In other words, each voter is given a weight equal to the natural logarithm of the odds that his judgment is correct. For example, if P_i is 0.75 (voter *i* is right 75 percent of the time) then the weight for his vote is ln(0.75/0.25) = ln(3) = 1.096. Note that a voter who is right exactly half the time receives a weight of ln(0.5/0.5) = ln(1) = 0, and a voter who is right less than half the time receives a negative weight.

Self-Interested Advocates

When voters are regarded as self-interested advocates, a different rationale for unequal predetermined weights for votes arises. If the collectivity is seeking the utilitarian outcome, that is, the outcome that yields the greatest net benefit to all voters, and if there is reason to believe that some voters, whatever their preferences, are affected more intensely than others, then the best statistical indicator of the outcome that provides the greatest total benefit is one that weights each voter's vote in proportion to the expected magnitude of the impact on him. For example, if a university was thinking about introducing an alternative calendar and wanted to let the faculty and students vote on the matter, it might be sensible to weight each vote according to an estimate of how much time that individual would spend under the new schedule (or, to be more precise, the expected present value of the amount of time under the new schedule).

For any actual example of unequal weights, it is difficult to be sure which rationale, if either, applies. The weighting of stockholders' votes according to the number of shares owned could be done because those who have more money invested have a greater incentive to become informed and might therefore be better judges. It could be because in the aggregation of the not-always-coinciding

interests of different stockholders it makes sense to weight more heavily the interests of those who have more money invested. Or it could be simply because if such weights were not used, individuals would have incentives to sprinkle bits of their holdings among friends and relatives to increase their influence, and such activity would in general be unproductive.

Summary

The historical origins of majority rule are obscure. Majority rule may have been an improvement upon a practice of simply acting when it seemed that the best proposal had been made, or it may have evolved from a practice of using coercion to achieve unanimous agreement. A probabilistic defense of majority rule can be made on the utilitarian basis that the majority is more likely than the minority to receive the higher total utility from having their way. Alternatively, if voters are viewed as disinterested judges, it can be argued that the majority is more likely than the minority to discern the better option. From the perspective of formal logic, majority rule can be defended as the only vote-processing rule with the properties of universal domain, anonymity, neutrality and positive responsiveness (May's theorem). Alternatively, positive responsiveness can be replaced by the combination of monotonicity and resolvability.

Defenses of *weighted* majority rule arise from relaxing the assumption in the probabilistic rationales that there is no basis for treating voters differently. If some voters are affected more intensely than others, then the utilitarian response is to weight their votes more heavily. Similarly, if some voters are better judges than others, their votes should be weighted more heavily to get the best estimate of the better course of action.

The next chapter explores the problem of cycles, which arises when there are more than two options among which a choice must be made.

Notes

1 J.A.O. Larsen, "The Origin and Significance of the Counting of Votes," *Classical Philology* 44 (1949): 164–81.
2 Ibid., p. 166.
3 See John G. Heinberg, "Theories of Majority Rule," *American Political Science Review* 26 (1932): 452–69.
4 Marie Jean Antoine Nicolas Caritat, Marquis de Condorcet, *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix* (Paris, 1785).
5 K.O. May, "A Set of Independent Necessary and Sufficient Conditions for Simple Majority Decision," *Econometrica* 20 (1952): 680–84.

6 For a fuller development of the ideas in this section, see T.N. Tideman, "A Majority-Rule Characterization with Multiple Extensions," *Social Choice and Welfare* 3 (1986): 17–30.

7 Monotonicity is sometimes called non-negative responsiveness.

Chapter 9

Voting Cycles

The Complications Introduced by Cycles

When more than two options are under consideration, cycles in majority rule are possible. The simplest example of cycles in majority rule occurs with three options and three voters, as in Example 9.1 below. Each column in Example 9.1 represents the order in which a voter ranks the three available options, *x*, *y* and *z*, with the most highly preferred option at the top.

Example 9.1 An example of cycles under majority rule

x	*y*	*z*
y	*z*	*x*
z	*x*	*y*

When majority rule is applied to pairs of options in Example 9.1, *x* beats *y* by two votes to one, *y* beats *z* by two votes to one, and *z* beats *x* by two votes to one. Each option beats one of the other two options by majority rule, while losing to the remaining option.

Suppose that a group attempted to choose among options *x*, *y* and *z* in Example 9.1, using a rule that pairs of options would be evaluated by majority rule until an option was found that could beat all others in paired comparisons. Suppose further that in each comparison, each voter always voted for the option that he or she ranked higher. Then the process of comparing would cycle endlessly among the three options. Thus the first lesson for voting over more than two options is that the rule, "Compare options in pairs until an option is found that can beat all others," is not a vote-processing rule that provides a result in all circumstances, because it is possible that there is no option that beats all others in paired comparisons.

One way of dealing with cycles under majority rule is to eliminate from further consideration any option that is defeated in a comparison with another option. Then, assuming that voters always vote in paired comparisons for the option of the two under consideration that they prefer, the outcome will depend on the order in which the options are considered. When three options are considered, whichever one is left out of the initial comparison will be the eventual winner. This can be

seen from the fact that each option wins one comparison and loses the other, so that the option that wins the first comparison will lose the second.

For cycles over more than three options there is, for each option, a set of options that beat it. The first option to have all the options that can beat it eliminated will necessarily be the winner. For a cycle in which x_1 beats x_2, which beats x_3, and so on until x_n, which beats x_1, x_i can be made the winner by having the first comparison between x_{i-2} and x_{i-1}, the next between the winner of that contest (namely x_{i-2}) and x_{i-3}, and so until $x_2, x_1, x_n, x_{n-1}, x_{n-1}, \ldots$, and finally x_i.

The dependence of the result on the order in which the options are compared gives crucial importance to the process by which this order is selected. If a random process is used to select the order, the outcome will depend on how the random process turns out. If there is a chairman who determines the order in which the options are considered, the outcome will depend on what he or she chooses. Since a chairman who knows the preferences of the members of the collectivity will be able to secure the selection of his choice, assuming that the members of the collectivity always vote for the options they prefer, such a practice is tantamount to allowing the chairman to select the outcome by the mode of authority.

What has been said up to this point assumes that each voter votes in each comparison for the option that he or she favors. Such voting is called *sincere* voting. Voting that is not sincere is called *strategic* voting. It can be seen that the assumption that voters will vote sincerely is not a plausible assumption when the number of voters is small and voters know one another's preferences. Return to Example 9.1 and suppose that the chairman is Voter 3. He announces that first options x and y will be compared, and then the winner of this contest will be compared with z. If the voters know each other's preferences, then Voter 1, who favors x, will be able to predict that though x will beat y in the first round, it will lose to z in the second, and he will wind up with his last choice. Foreseeing this result, he might reasonably decide to vote for y in the first round, resulting in y winning the first round and then beating z in the second, so that Voter 1 will at least wind up with his second choice rather than his third. The chairman now winds up with his third choice, despite having set the agenda.

It is possible that Voter 1 could do even better. After Voter 3 announces the agenda, Voter 1 tells Voter 3 his strategic plan and then offers to make a deal. Voter 1 will vote for x instead of y in the first round if Voter 3 will vote for x instead of z in the second round. That way Voter 3 will get his second choice rather than his third, which he would get if Voter 1 carried through with his strategy. If Voter 3 agrees and does as he promises, Voter 1 gets his first choice instead of the third choice that sincere voting by everyone would have gotten him.

If Voter 2 becomes aware of the conversation between Voters 1 and 3, he may say to Voter 3, "You don't have to make a deal with Voter 1. If you turn him down, I will agree to vote for z in the second round." Then z wins. Or Voter 1 may break in and say to Voter 2, "If you refrain from agreeing with him, I will

support *y* in both the first round and the second." Then *y* wins. Thus any outcome is possible, depending on which coalition is formed.

The agenda need not be structured as a sequence of paired comparisons, with the loser in each comparison eliminated. An alternative way of structuring the decision would be to consider proposals one by one, with the considered proposal chosen if it received a majority, and otherwise eliminated. Assume that sincere voting now means voting for the proposal under consideration only if it is one's first choice among those not yet eliminated. If voters vote sincerely, the chairman's strategy is to consider first the option that beats his first choice. Since it is not the first choice of a majority, it will be eliminated and the chairman's first choice will be chosen. Thus a chairman who favored *z* would have *y* considered first. And Voter 1 could defeat the chairman's strategy by voting for *y*.

The general conclusion to be drawn from these possibilities is that when there are cycles, the ability to set the agenda is valuable if voters vote sincerely, but strategic voting and the formation of coalitions can nullify the power of the agenda setter and yield any outcome in the cycle. The existence of cycles thus negates some of what is achieved by agreeing on a procedure instead of seeking to agree on an outcome. The outcome is not a simple function of the procedure, but instead depends on what coalition agreements are reached, and therefore on the political skills of the participants.

Conditions that Preclude or Promote Voting Cycles

One example shows that voting cycles can happen. But are they *likely* to happen? Insight into this question comes from examining conditions that preclude or promote voting cycles. More relevant than the question of whether there are cycles is the question of whether there is a "majority-rule equilibrium." A majority-rule equilibrium exists if there is an option that does not lose to any other option by majority rule. If there are no cycles, then there is a majority-rule equilibrium, but if there are cycles, there could still be a majority-rule equilibrium if no cycle includes the equilibrium option. It can also happen that there are ties involving the equilibrium option, so that, although the equilibrium option does not defeat all others, neither is it defeated.

Suppose that the options under consideration have positions on a one-dimensional continuum, as in Figure 9.1. Then a voter's preferences are said to be "single-peaked" if, for any pair of options that are on the same side of the voter's first choice, the voter prefers the option that is closer to his first choice. The preferences among seven options shown in Figure 9.1 are single-peaked. If every voter's preferences for options on a one-dimensional continuum are single-peaked, then there is a majority-rule equilibrium when voters vote sincerely.[1]

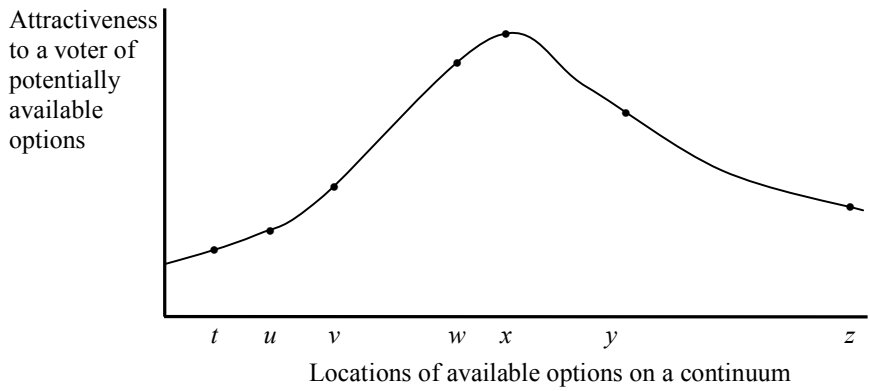


Figure 9.1 An example of single-peaked preferences

The proof is as follows: Let n be the number of voters. For each option, p, define $f(p)$ as the number of voters who find p the most attractive option. (For simplicity, we exclude the possibility that a voter finds two or more of the available options equally most attractive.) Define $g(p)$ as $f(p)$ plus the sum of $f(p)$ for all options, q, that are to the left of p. For example, if the numbers of voters who find options t through z most attractive are 4, 2, 1, 0, 2, 3, and 2, then $f(p)$ and $g(p)$ are as shown in Table 9.1.

Table 9.1 Values of $f(p)$ and $g(p)$ for the example in the text

Option	t	u	v	w	x	y	z
$f(p)$	4	2	1	0	2	3	2
$g(p)$	4	6	7	7	9	12	14

Identify the first option from the left such that $g(p)$ is at least half of n. Call that option r. Identify the first option from the left such that $g(p)$ is more than half of n. Call that option s.

For the example, r is option v and s is option x. No other option beats r or s in paired comparisons. The reason no option beats r is that when r is compared with an option to its right, r is preferred by the voters whose first choice is r, and, since preferences are single-peaked, by all the voters whose preferences are to the left of r. From the definition of r, these constitute at least half the voters. When r is compared with an option to its left, r is preferred by the voters whose first choice is r, and, since preferences are single-peaked, by all the voters whose preferences are to the right of r. From the definition of r, these constitute more than half the voters. Thus no option can obtain more than half the votes in a comparison with r. The proof that no option beats s is the same, except that "left" and "right" are

interchanged. Since nothing beats *r* or *s*, they are majority-rule equilibria. In fact, by a similar proof, any option between *r* and *s* is a majority-rule equilibrium. Thus *v*, *w*, and *x* are all majority-rule equilibria in the example. If *r* and *s* are the same option, this option is called the preference of the median voter and is the unique majority-rule equilibrium. By extension, one can refer to *r* and *s* together as the preferences of the median voters. Thus the general conclusion about cycles and single-peaked preferences can be stated as, "If all voters' preferences in a continuum of one dimension are single-peaked, then, under sincere voting, there is a majority-rule equilibrium (or there are majority-rule equilibria) at the preference(s) of the median voter(s)."

The ability of single-peaked preferences in one dimension to preclude cycles raises the question of how likely single-peaked preferences are. Economists often assume that preferences are "convex," meaning that if a person finds two options, *x* and *y*, equally attractive, then he will find anything between *x* and *y* on the straight line connecting them at least as attractive as *x* or *y*. For options that lie in a single dimension, convexity and single-peakedness are equivalent. The convexity of preferences is connected to the notion of "diminishing marginal rates of substitution," meaning that the more of any one thing you have, the more of it you are willing to sacrifice to obtain one more unit of something else. Economists observe that people's preferences generally have this characteristic, which increases the plausibility of single-peakedness.

The preferences that many people have for temperature in tea provide one example of preferences *not* being convex: They regard tepid tea to be less attractive than both hot tea and cold tea. It has been pointed out that there are circumstances in which it is reasonable for people's preferences for political programs to be non-convex. Suppose that people will send their children to private schools if the public schools are not good enough to suit them, and that they rank alternative levels of funding for schools purely in terms of net benefits for themselves. In that case some people may find high and low quality schools both more attractive than moderate quality schools because with moderate quality schools their taxes will be higher than with low quality schools, but the quality of the schools will not be high enough for them to want to send their own children.[2] Despite such examples, it seems reasonable to regard convex preferences, and preferences that are single-peaked in one dimension, to be the norm, to which there will be occasional exceptions.

The ability of single-peaked preferences to preclude cycles is limited to instances when the options lie in a continuum of *one* dimension. It is not entirely clear what single-peakedness ought to mean when the concept is extended to two dimensions, so take a case of the most regular possible preferences, where between any two options, any voter always prefers the one that is closer to his first choice. In Figure 9.2, *x*, *y*, and *z* are three options that are under consideration, and *A*, *B*,

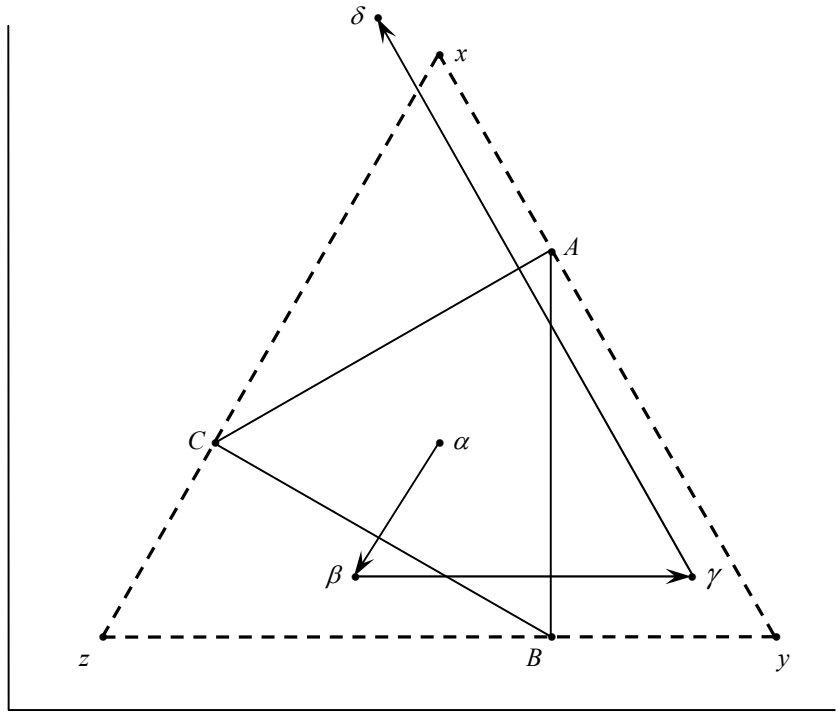


Figure 9.2 Options in a plane, preferences determined by distance

and *C* are the locations that are most preferred by three voters, who will be given the same names. Note that voter *A* ranks the options *xyz*, voter *B* ranks them *yzx*, and voter *C* ranks them *zxy*, so that the cycling of Example 9.1 is produced.

It is interesting that if voters always vote sincerely, a diabolical agenda setter with access to all potential outcomes could move the outcome to any point in the space of outcomes.[3] Thus, for example, starting at α, which is equidistant from the ideal points of the three voters, β is preferred to α by the majority consisting of voters *B* and *C*, γ is preferred to β by the majority consisting of voters *A* and *B*, and δ is preferred to γ by the majority consisting of voters *A* and *C*.

Is cycling *likely* when options are points in a continuum of two or more dimensions? Suppose that all points in the continuum are options that could be chosen. Consider a potential equilibrium point, *x*. For each voter, *i*, there is an indifference curve D_i through *x*, of options that are just as good as *x* in the estimate of that voter. On one side of D_i are options that *i* regards as better than *x*, while on the other side of D_i are options that the voter regards as worse than *x*. For *x* to be a majority-rule equilibrium, it must be true that for *any* movement away from *x*, the voters who thereby move to the preferred sides of their indifference curves through *x* are no more than half of the voters. It seems rather unlikely that the ideal points

would be so smoothly distributed that any point would possess this property. Thus it must be expected that for choices over a continuum of more than one dimension, there will generally be no majority-rule equilibrium.[4]

Nevertheless, cycles will not necessarily be a practical problem. If, as seems plausible in many circumstances, there is a tendency for voters' first choices to be approximately symmetrically distributed, and if the agenda involves the consideration of only a few options that are not exceedingly close to one another, then the prospect of finding a majority-rule equilibrium among the options on the agenda is considerably improved.[5]

A circumstance in which a majority-rule equilibrium *cannot* be expected is when voters are selfish and the options under consideration are purely redistributive. From any starting point it is always possible to propose an alternative that takes something from some minority of the voters and distributes it among a majority, so that a majority would prefer the alternative to the starting point.

The severity of the cycles that can occur when an attempt is made to decide distributions can be appreciated from the following example: Suppose that 100 persons are attempting to distribute among themselves 100 positions of varying attractiveness. The positions might be plots of land, or dormitory rooms, or work assignments. There would be something to be said for auctioning off the positions and distributing the proceeds of the auction, but suppose that is not going to happen. Barring an auction, there would be something to be said for an allocation by random process, but that is excluded too. The group is going to use majority rule to decide the order in which individuals will be allowed to choose positions. They realize that there are $100 \cdot 99 \cdot 98 \cdot \ldots \cdot 3 \cdot 2$, that is, 100! different possible orders in which the individuals might choose, so since they cannot possibly consider all of these, they decide to confine their attention to orders of choosing that start somewhere in the alphabet, proceed to the end, move to the beginning of the alphabet and continue alphabetically until everyone has chosen. The only thing to be decided is where in the alphabet they will start. Each proposal to start one person later in the alphabet will be opposed by the person who thereby moves from first to last, and favored by the other 99 voters. Thus there will be a cycle among the 100 possibilities, with each step in the cycle approved by a margin of 99 to 1. Thus it would be inappropriate to think of cycles as a problem that arises simply because a bare majority can pass a measure. No matter how high a majority is needed to pass a measure, cycles are possible.[6]

The Frequency of Cycles in Real Voting Situations

The previous section established that majority-rule cycles can occur, but that one should expect them not to occur very often among real options, except when the

options are purely redistributive. But how often do cycles among real options actually occur?

The existence of logrolling in legislative bodies is evidence of real cycles. If there are two projects, x and y, that are both defeated when considered separately, while a proposal to undertake both projects secures a majority, then there is a cycle among the possibilities, "undertake x and y," "undertake just x," and "undertake neither project."[7]

It is the potential for redistribution through public projects that makes voting cycles with respect to them inevitable. Cycles with respect to candidates for offices require separate consideration. Data relevant to exploring the frequency of cycles among candidates are scarce because the voting procedures that are most commonly used ask voters only to vote for one candidate. The voters' full rankings of candidates are not sought, and therefore cycles can exist without providing evidence of their existence.

However, some voting procedures that are actually used do ask voters to provide their full rankings of candidates. In particular, the Single Transferable Vote (STV), which is discussed in detail in Chapter 15, asks voters to provide full rankings of candidates. By examining ballots from STV elections it is possible to obtain evidence concerning the empirical frequency of cycles.

The Electoral Reform Society of Great Britain and Ireland (hereafter the ERS) promotes the use of STV and, through its Ballot Services division, oversees both elections by STV and other types of elections. Ballots from 84 ERS-supervised elections in which voters were asked to rank candidates were put onto computer tape. In 78 of these elections STV was employed. The other 6 elections were conducted by "multi-vote plurality," where each voter is given as many votes as there are positions to be filled. Voters were asked to rank the candidates in these elections because both delegates and alternates were being elected, and those who were defeated as delegates could be elected as alternates. Three elections that had been gathered by I.D. Hill were added to the 84 elections from the ERS.[8] A summary of some characteristics of these 87 elections is provided in Table 9.2.

When data from STV elections are examined, it is immediately apparent that, although voters are asked to rank all candidates, they typically leave a substantial fraction of the candidates unranked. In the data used for this study, the average proportion of candidates ranked (weighting all elections equally, irrespective of the number of voters) was 0.604. Table 9.3 shows the average proportion of the candidates ranked by voters, sorted by the number of candidates in an election and the number to be elected. Table 9.3 shows that the proportion of candidates that voters rank falls with the number of candidates and rises with the number to be elected.

An implication of the limited number of candidates that voters rank is that when these data are used to inquire into the probability of a cycle, much of the desired data will be missing. Candidates who are not ranked must be treated as tied

Table 9.2 A summary of characteristics of 87 elections employed in investigating the probability of a voting cycle

Number of candidates		Number elected		Number of voters [a]	
Range	Number of cases	Range	Number of cases	Range	Number of cases
3 – 4	25	1	20	9	1
5 – 6	14	2	25	30 – 59	6
7 – 9	17	3 – 4	22	60 – 99	12
10 – 14	19	5 – 8	13	100 – 199	19
15 – 19	7	9 – 12	7	200 – 499	22
20 – 29	5			500 – 999	24
				1000 – 3500	3

[a] For 29 of the 87 elections, the number of voters represents a sample of the actual election.

at the bottoms of rankings. But it is still possible to observe cycles. A voter is regarded as ranking candidate x ahead of candidate y either if x is given a higher rank than y on the voter's ballot or if x is mentioned and y is not. Candidate x is said to have a majority over candidate y if the number of voters who rank x higher than y is greater than the number who rank y higher than x. A cycle exists if there is a sequence of candidates (z, y, x, …, i) such that z has a majority over y, y has a majority over x, …, and i has a majority over z. A "semi-cycle" will be said to exist if there is a sequence of candidates that satisfies the above condition when "has a majority over" is replaced by "has a majority over or is tied with," and at least one relationship in the sequence is a majority. A "cyclic anomaly" means either a cycle or a semi-cycle. In addition to cyclic anomalies, it will be interesting to take note of "simple ties," which occur when there is a tie that is not part of a semi-cycle.

Table 9.3 Average proportion of candidates ranked, sorted by number of candidates and number elected (all elections weighted equally)

Number of candidates	Number elected				
	1	2	3 – 4	5 – 8	9 – 12
3 – 4	.749	.783	.880		
5 – 6	.513	.604	.724		
7 – 9	.535	.476	.569	.799	
10 – 14	.378	.383	.473	.481	.639
15 – 19		.312	.390	.480	
20 – 29		.234			.478

An election contains no cyclic anomalies and no ties if and only if there is a way of naming the candidates, $x_1, x_2, \ldots, x_n$ such that for any $i, j \leq n$, if $i < j$ then x_i has a positive majority over x_j. One way of discovering whether this condition is satisfied is to count the number of other candidates over whom each candidate has a majority and then order the candidates by the number of candidates they beat. An election among n candidates has no cyclic anomalies and no simple ties if and only if there is one candidate who beats $n - 1$ candidates, one who beats $n - 2$ candidates, and so on down to one candidate who beats one other candidate and one who beats none. An ordering of the candidates such that each candidate beats by majority rule all of the candidates that follow him will be referred to as a transitive ordering of the candidates.

A computer program was used to seek a transitive ordering of the candidates in the 87 elections in the sample. For 61 elections the program found a transitive ordering. For the other 26 elections, the program placed the candidates in the order that came closest to being transitive, where closeness is measured by a criterion suggested by Condorcet, namely that the sum of the majorities of pairs where a lower candidate beats a higher candidate be as low as possible.[9]

To employ the Condorcet criterion, one calculates a "matrix of majorities," that is, a matrix in which component (i, j) is the majority of candidate i over candidate j, the difference between the number of voters who rank i ahead of j and the number who rank j ahead of i. (The diagonal components of the matrix of majorities are zeros.) Table 9.4 shows the matrix of majorities for one election in the sample. For this election (election 57 in the data), there is a cycle among candidates 8, 2, 7 and 3, but candidate 4 is nevertheless a majority-rule equilibrium.

If there is a transitive ordering of the candidates, then for that ordering the upper triangle of the matrix of majorities consists entirely of positive numbers. If there is a majority-rule equilibrium, then the row of the matrix of majorities representing that equilibrium will consists entirely of non-negative integers. Thus

Table 9.4 An example of a matrix of majorities

	4	5	9	8	2	7	3	1	6
4	0	30	103	127	102	132	121	136	128
5	–30	0	122	154	128	155	142	164	156
9	–103	–122	0	26	32	31	28	54	52
8	–127	–154	–26	0	3	3	–4	18	12
2	–102	–128	–32	–3	0	16	21	11	36
7	–132	–155	–31	–3	–16	0	5	13	7
3	–121	–142	–28	4	–21	–5	0	7	15
1	–136	–164	–54	–18	–11	–13	–7	0	8
6	–128	–156	–52	–12	–36	–7	–15	–8	0

one way of exploring the probability of an election outcome that is free of cyclic anomalies and simple ties is to ask, "What is the probability that, for the ordering that is best by the Condorcet criterion, every component of the upper triangle of the matrix of majorities will be positive?" Similarly, to investigate the probability of a majority-rule equilibrium, one can ask, "What is the probability that, for the ordering that is best by the Condorcet criterion, every component of the top row of the upper triangle of the matrix of majorities will be positive?" To investigate the probability of a simple tie, one can ask, "What is the probability that a component of the first off-diagonal of the upper triangle of the matrix of majorities will be zero?" To investigate these questions, one needs a theory of the determinants of the magnitudes of the components of the matrix of majorities.

Toward that end, note that the matrices of majorities of the 87 elections in the sample reveal five patterns:

1. Elections with few candidates almost always have transitive orderings. Elections with many candidates virtually never have transitive orderings. Figure 9.3 shows how the frequency of cyclic anomalies varies with the number of candidates, for elections in the sample. The basic source of the pattern in Figure 9.3 is simply that there are more opportunities for cyclic anomalies in elections with many candidates. There is no pattern to the average frequency with which components of matrices for elections with different numbers of candidates are anomalous.

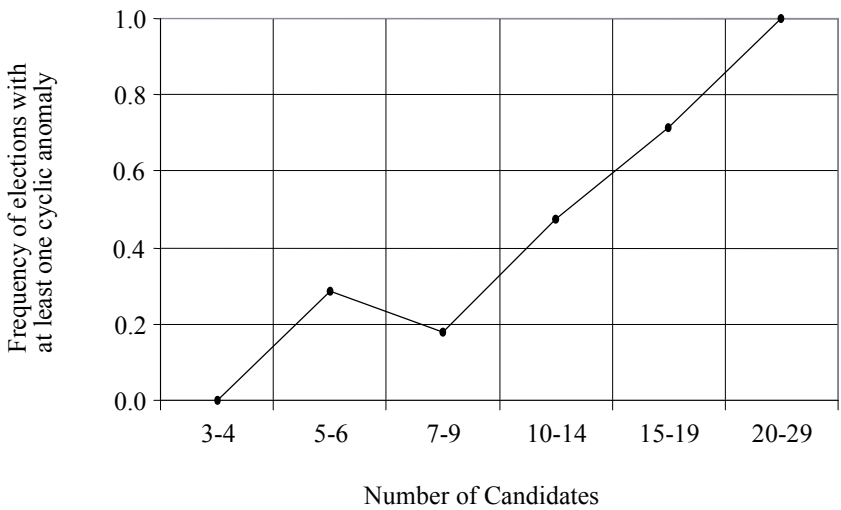


Figure 9.3 Frequency of elections containing at least one cyclic anomaly as a function of the number of candidates in the election

2. Pairs ranked by few voters are more likely to generate cyclic anomalies than pairs ranked by many voters. Figure 9.4 shows how the frequency of cyclic anomalies varies with the number of voters who rank a pair, for elections in the sample.

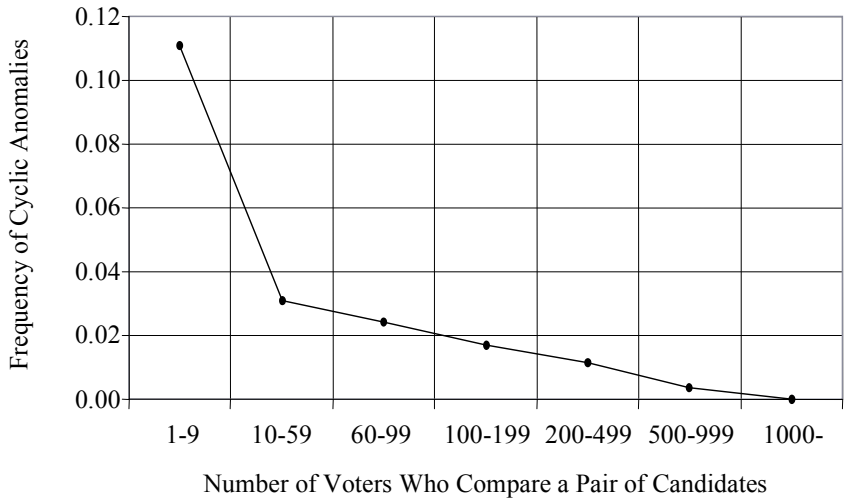


Figure 9.4 Frequency of cyclic anomalies in components of matrices of majorities as function of the number of voters who compare a pair of candidates

3. Even allowing for the fact that cyclic anomalies involve pairs that are ranked by few voters, the majorities involved in cyclic anomalies are extremely small. The average positive component of a matrix of majorities (121.3) is 53.7 times the size of the average negative component (2.26). Figure 9.5 shows the distribution of majorities in cyclic anomalies.

4. Anomalous components of the matrix of majorities tend to be located near the main diagonal. Figure 9.6 shows the frequency of cyclic anomalies as a function of the relative distance from the main diagonal. "Relative distance" is computed as a ratio in which the numerator is the difference between the ranks of the candidates that generate the anomalous component and the denominator is the number of candidates plus one. The virtue of this measure of relative distance is that if one takes three candidates at random from any number of candidates greater than two, the expected value of the relative distance of the component comparing the first in rank with the third in rank is 0.5.[10] Thus the expected value of the relative distance assigned to an anomaly is independent of the number of other candidates in an election.

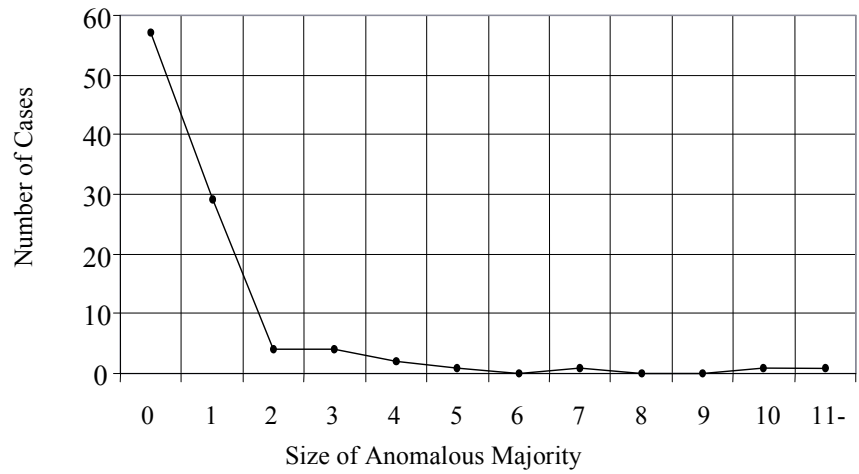


Figure 9.5 Distribution of majorities for components of matrices of majorities representing cyclic anomalies

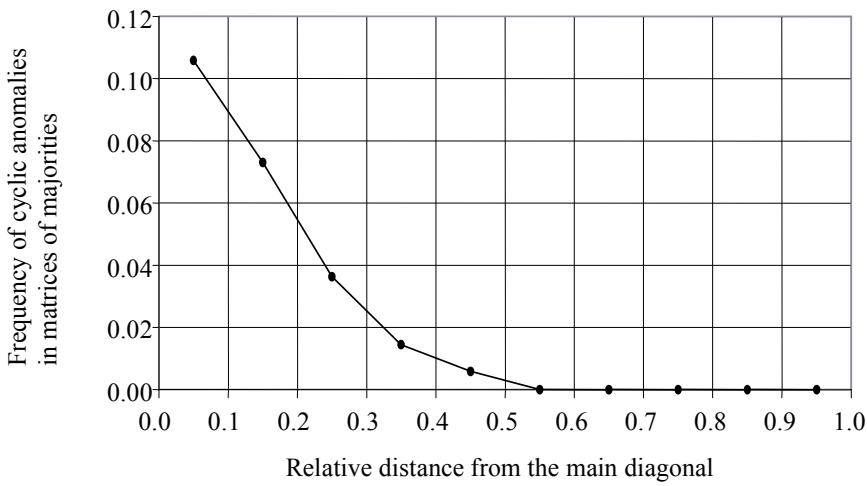
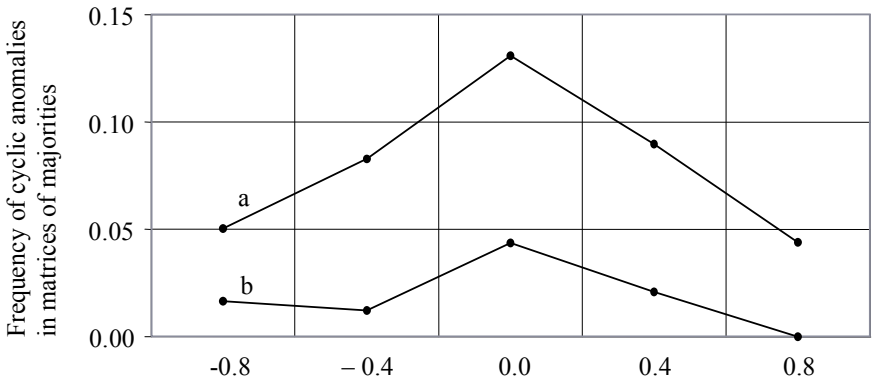


Figure 9.6 Frequency of cyclic anomalies in components of matrices of majorities as a function of relative distance from the main diagonal

5. On any given diagonal, anomalous components of the matrix of majorities tend to lie near the center of the diagonal. This is illustrated in Figure 9.7, which shows the frequency of cyclic anomalies in components of matrices of majorities as a function of relative distance from the center of a diagonal, for two ranges of relative distance from the main diagonal.



a: relative distance from the main diagonal is 0.03 to 0.09
b: relative distance from the main diagonal is 0.09 to 0.24

Figure 9.7 Frequency of cyclic anomalies in matrices of majorities as a function of relative distance from the center of a diagonal

To recapitulate, the following five patterns have been observed: Cyclic anomalies tend to occur in elections with many candidates and in elections with few voters. The anomalies tend to involve small majorities. And anomalies tend to occur close to the main diagonal of the matrix of majorities and close to the center of the diagonal on which they occur. The model that follows accounts for all five of these patterns.

Informal presentation of the model

Assume that the Condorcet criterion ranks candidates in terms of "over-all attractiveness," which has a roughly normal distribution in the population from which candidates are drawn. Then the greater the number of candidates, the greater will be the proximity, in over-all attractiveness, of any candidate to the candidates who are ranked near him. Cyclic anomalies tend to occur when the agenda contains candidates who are similar in over-all attractiveness. Thus cyclic

anomalies are more prevalent when the number of candidates is greater, because then candidates will be closer in over-all attractiveness to each other. Candidates tend to be closer in over-all attractiveness to the candidates who are ranked near them. Thus cyclic anomalies are more likely to occur on diagonals near the main diagonal (though not on the first off diagonal, except ties, because there a reversal causes a change in ranking rather than an anomaly). And because the distribution of over-all attractiveness is roughly normal, candidates in the center of the ranking are more similar to their neighbors than those at the extremes. Thus anomalies are more likely to occur toward the center of any diagonal on which they occur.

Given a pair of candidates, i and j, with i ahead of j in the over-all ranking, the probability, z, that a voter will rank candidate i below candidate j, can be treated as a random variable with a distribution that depends on the number of candidates and on the location of component (i, j) in the matrix of majorities. Table 9.5 shows how the average value of z varied with the location of a paired comparison in the matrix of majorities, for the 87 elections for which rankings were collected.

It is reasonable to suppose that in the limit as the number of candidates approaches infinity, the mean value of z for a pair of neighboring candidates approaches 1/2, because neighboring candidates then become indistinguishable. In terms of Table 9.5, this means that as the number of candidates approaches infinity, the values along the lower-left edge of the table should approach 1/2.

Table 9.5 Average value of z as a function of location in the matrix of majorities

Relative distance from the center of a diagonal (y)	Relative distance from the main diagonal (x)							
1	0							
0.420	0.343	0.312	0.267	0.233	0.205	0.194	0.127	
0.403	0.334	0.300	0.290	0.247	0.208	0.149		
0.413	0.374	0.341	0.300	0.281	0.223	0.157		
0.419	0.404	0.356	0.313	0.275	0.209			
0.440	0.409	0.378	0.328	0.280	0.181			
0.438	0.434	0.367	0.322	0.256				
0.449	0.409	0.380	0.313	0.190				
0.450	0.409	0.358	0.269					
0	0.449	0.385	0.347	0.242				1
	0.441	0.387	0.301					
	0.439	0.381	0.307					
	0.437	0.339						
	0.420	0.309						
	0.395							
	0.389							
	1							

The mean value of z should decline as one moves away from the main diagonal (as one moves to the upper-right in Table 9.5), since proximity to the main diagonal is a measure of how close to each other candidate are. The mean value of z should decline more slowly if one starts at the center of the main diagonal than if one starts away from the center, since adjacent candidates at the center of the main diagonal are closer to each other than adjacent candidates away from the center.

It is possible for the random process that generates z to yield a value that is greater than 1/2, although the mean value yielded by the process would be less than 1/2. A value of z greater than 1/2, if it occurred beyond the first off-diagonal, would lead to a cyclic anomaly with a probability that approached 1 as the number of voters increased. But a cyclic anomaly can also arise when z is less than 1/2, if it should happen by chance that more than half of the actual voters rank i below j, even though that would not happen in the limit as the number of voters approached infinity. If z is drawn from a distribution almost all of which is significantly less than 1/2, then cyclic anomalies will be rare. And when they do occur, they will tend to involve small majorities. For values of z that are close to but less than 1/2, anomalies will be more common when the number of voters is small than when it is large. Thus the model accounts for the five observed regularities with respect to cyclic anomalies. The next steps are to make the model analytically precise and to estimate its parameters.

Formal Development of the Model

For any paired comparison in any election, the probability that a voter chosen at random will rank candidate i below candidate j, z_{ij}, is assumed to be a random variable that is distributed according to a beta distribution. For simplicity of notation, the subscripts of z will be suppressed hereafter. The assumption of a beta distribution implies that the probability that z will lie in a narrow interval around some number z_0 between 0 and 1 is equal, in the limit as the width of the interval approaches zero, to the width of the interval multiplied by

$$\frac{1}{B(p,q)} {z_0}^{p-1}(1-z_0)^{q-1}, \tag{9.1}$$

where $B(p, q)$ is the normalizing constant for the beta distribution,

$$\frac{(p-1)!(q-1)!}{(p+q-1)!}, \tag{9.2}$$

that insures that the integral of the density function is 1, and p and q are positive numbers that must be estimated.

For the purpose of modeling cyclic anomalies, it will be convenient to estimate p and q indirectly, through the parameters r and n, defined by $r = p/(p + q)$ and $n = p + q$. The mean of the beta density function is r and its variance is $r(1 – r)/(n + 1)$. For values of n greater than 2, the beta density function has a shape similar to the normal density function, except that the distribution is confined to the interval between 0 and 1, with a mode that can occur anywhere in the interval. Figure 9.8 shows an example of the beta density function.[11]

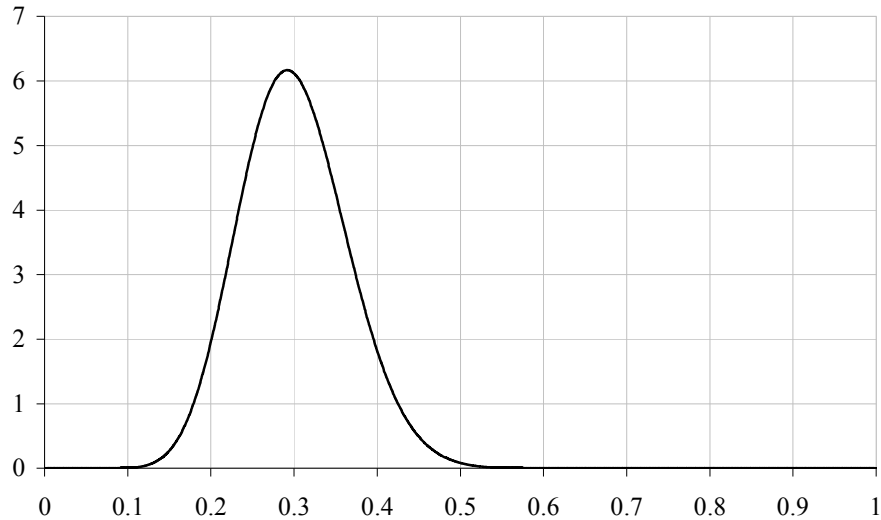


Figure 9.8 An example of the beta density function, with $r = 0.3$ and $n = 50$

Roughly speaking, the parameter r determines the location of the mode of the distribution (the greater the value of r, the farther to the right the mode is), while the parameter n determines how narrow the distribution is (the greater the value of n, the narrower the distribution is).

For estimating the probability of an anomaly in the comparison of a pair of candidates, the parameter r is equivalent to the mean of the distribution from which z, the probability of voting for a lower-ranked candidate, is drawn. Therefore r is assumed to be a function of the location of the comparison in the matrix of majorities. The comparison of the candidates ranked i and j determines components (i, j) and (j, i) of the matrix of majorities. Without loss of generality, it will be assumed that i is the higher ranked candidate, where the highest rank is 1. The comparison of i and j is assigned x and y coordinates in a coordinate system that is oriented like Table 9.5, with a range of (–1, 1) for y and a range of (0, 1) for x. The formulas for the coordinates are:

$$x = (j – i)/(k + 1); \tag{9.3}$$

$$y = 1 - (j + i)/(k + 1), \tag{9.4}$$

where k is the number of candidates.

The parameter r is assumed to vary according to a function of the form

$$r = 0.5\, e^{a_1 x + a_2 x^2 + a_3 x y^2}. \tag{9.5}$$

The coefficient of 0.5 at the beginning of the equation arises from the fact that in the limit as the number of candidates approaches infinity, the x coordinate for a pair of neighboring candidates approaches 0, and the probability of voting for the lower of two approaches 0.5. When x is 0, the exponent for e is 0, and the exponential term becomes 1.

The complicated exponent for e was chosen to reflect several theoretical considerations and observations about the data. The first two terms of the exponent, $a_2 x$ and $a_3 x^2$, reflect the manner in which the average probability of voting for the lower-ranked candidate varies for pairs of candidates that are equidistant from the middle of the over-all ranking, one above and one below, that is, pairs of candidates for which y is 0. The coefficient on the first term reflects the rate at which r decreases when x is zero, and the need for the second term reflects the fact that as x increases, the rate at which r falls increases at first and then decreases. This is a natural consequence of candidates being closer in over-all attractiveness at the center of the over-all ranking than at the ends.

The third term of the exponent, $a_3 x y^2$, reflects the fact that as x increases, r falls more rapidly at the ends of the over-all ranking than in the middle.

Next, the parameter n must be specified as a function of parameters that are to be estimated. To observe the variations in n, Table 9.6 was constructed. This shows how the value of n implied by the data varies in xy-space, where the implied value of n, derived from the formula for the variance of the beta distribution, is given by

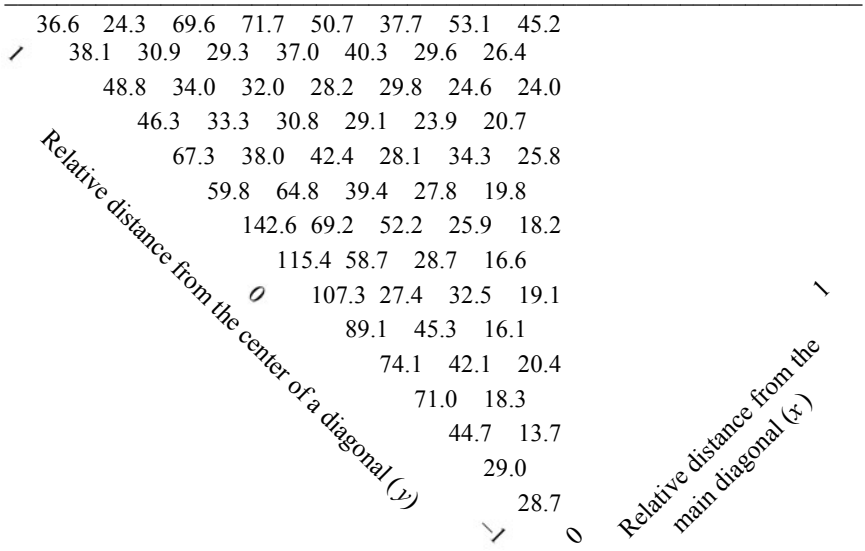
$$\frac{(\text{mean of } r)(1 - \text{mean of } r)}{\text{variance of } r} - 1. \tag{9.6}$$

Table 9.6 has the same dimensions and orientation as Table 9.5, which shows variations in the implied mean value of r.

It is assumed that n varies according to the equation

$$n = [\, r\,(.5 - r)\,]^{a_4}. \tag{9.7}$$

Table 9.6 Implied value of *n* as a function of location in the matrix of majorities



This form reflects the idea that *n* should approach infinity (the distribution of *r* should become arbitrarily narrow) in the limit as *r* approaches either 0 or 0.5. Equation 9.7 has this shape for negative values of a_4.

Given a set of assumed values for parameters a_1 through a_4, the likelihood of an observed result of a paired comparison in which *v* voters voted for the candidate with rank *i* out of *k* and *w* voters voted for the candidate with rank *j* out of *k* can be computed in the following way. First compute *r* from equations 9.3 – 9.5 and *n* from equation 9.7, from which *p* and *q* can be computed. Then probability theory implies that the probability of observing *v* and *w* voters voting as they did is

$$\int_0^1 f(z)P(v,w|z)dz\,. \tag{9.8}$$

where *f* is the density function for *z* and $P(v, w|z)$ is the probability of observing *v* and *w*, given *z*. If *i* and *j* differ by more than 1 or if $w = v$, then this is

$$\int_0^1 \frac{1}{B(p,q)} z^{p-1}(1-z)^{q-1} z^w (1-z)^v \frac{(v+w)!}{v!\,w!} dz\,. \tag{9.9a}$$

But if *i* and *j* differ by only 1 and $w \neq v$, then it is

$$\int_0^1 \frac{1}{B(p,q)} z^{p-1}(1-z)^{q-1}[z^w(1-z)^v + z^v(1-z)^w]\frac{(v+w)!}{v!\,w!}dz\,. \qquad (9.9b)$$

The reason for the exception is that if, in the comparison of two adjacent candidates, the "lower" candidate should happen to get more votes than the "higher" candidate, the result is not an anomaly but rather an interchange in the positions of the two candidates.

Combining and rearranging terms in (9.9a) yields

$$\frac{1}{B(p,q)}\frac{(v+w)!}{v!\,w!}\int_0^1 z^{p-1+w}(1-z)^{q-1+v}dz\,. \qquad (9.10)$$

The expression that is integrated has the form of a beta density function, except that the normalizing constant is missing. Multiplying and dividing by this normalizing constant yields

$$\frac{B(p+w,q+v)}{B(p,q)}\frac{(v+w)!}{v!\,w!}\int_0^1 \frac{1}{B(p+w,q+v)} z^{p-1+w}(1-z)^{q-1+v}dz\,. \qquad (9.11)$$

Since the integral is a standardized density function, its value is 1, and 9.11 reduces to

$$\frac{B(p+w,q+v)}{B(p,q)}\frac{(v+w)!}{v!\,w!}. \qquad (9.12)$$

Substituting for the values of the normalizing constants, the likelihood of a vector of parameter values is

$$h = \frac{(p+w-1)!(q+v-1)!}{(p+q+v+w-1)!}\frac{(p+q-1)!}{(p-1)!(q-1)!}\frac{(v+w)!}{v!\,w!}. \qquad (9.13a)$$

A similar line of reasoning applied to (9.9b) yields

$$h = \frac{(p+w-1)!(q+v-1)! + (p+v-1)!(q+w-1)!}{(p+q+v+w-1)!}\frac{(p+q-1)!}{(p-1)!(q-1)!}\frac{(v+w)!}{v!\,w!}. \qquad (9.13b)$$

Because of the very small numbers that are involved in likelihoods, it is more convenient to work with the natural logarithm of the likelihood, which is maximized at the same parameter values as the likelihood itself. The log of the likelihood in terms of all the comparisons in an election is the sum of the logs of

the likelihoods of the comparisons taken separately. And the log of the likelihood for all the elections in the sample is the sum of the logs of the likelihoods of the elections considered separately. The parameter values that best explain the observed variations in the frequency of voting for a lower-ranked candidate are found by searching for the values that maximize this sum of logs of likelihoods.

A computer program was written to search for the parameter values that maximized the log of the likelihood. The values on which it settled are shown in Table 9.7.

Table 9.7 Estimated parameters of the determinants of the probability of a cycle

	a_1	a_2	a_3	a_4
Parameter estimate	– 0.532	– 0.789	– 2.486	– 1.281
Standard error [a]	0.028	0.055	0.010	0.011

[a] The standard errors are computed as the square roots of the diagonal components of the inverse of the matrix of second partial derivatives of the negative of the loglikelihood function.

The estimates of parameters a_1, a_{12} and a_3 imply that the mean of the distribution of z for any paired comparison is given by

$$r = 0.5\, e^{-0.532\, x - 0.789\, x^2 \; - 2.486\, xy^2}. \tag{9.14}$$

And the estimate of parameter a_4 implies that n, which determines the narrowness of the distribution of z is given by

$$n = [\, r\,(0.5 - r)\,]^{\,-\,1.281} \tag{9.15}$$

Using equations 9.14 and 9.15, it is possible to compute the values of r and n associated with any paired comparison. These values depend only on the number of candidates in an election and the positions in the over-all ranking of the candidates being compared. Table 9.8 shows calculations of r for the case of 10 candidates.

It is now possible to estimate the probability of a cycle or a tie for any paired comparison, if the number of voters who rank the pair is also known. This involves an application of the theory of Bernoulli trials, with a probability of success that is a random variable. For example, if there are 50 voters, then the probability of a tie between the first two of ten candidates is the probability of 25 "successes" and 25 "failures" in 50 independent trials where the probability of

Table 9.8 *r* as a function of *i* and *j* for the case of 10 candidates

					j				
	2	3	4	5	6	7	8	9	10
1	0.420	0.368	0.333	0.308	0.287	0.267	0.246	0.220	0.191
2		0.443	0.403	0.373	0.347	0.321	0.293	0.259	0.220
3			0.459	0.428	0.399	0.368	0.334	0.293	0.246
4				0.470	0.441	0.408	0.368	0.321	0.267
i 5					0.473	0.441	0.399	0.347	0.287
6						0.470	0.428	0.373	0.308
7							0.459	0.403	0.333
8								0.443	0.368
9									0.420

success in each trial is a random variable, z, that has a beta distribution with parameters of $r = 0.420$ (from Table 9.8) and $n = 77.3$ (from equation 9.15), implying that $p = 32.47$ and $q = 44.83$. Thus, from equation 9.13a, the probability of a tie for this case is estimated to be

$$\frac{56.47!68.83!}{126.3!}\frac{76.30!}{31.47!\,43.83!}\frac{50!}{25!\,25!} = 0.0586. \tag{9.16}$$

Similarly, the probability of a cycle among the top three of ten candidates comes from $r = 0.368$ and $n = 48.3$, implying $p = 17.77$ and $q = 30.53$. Then, using 9.13a again, the probability of a cycle is estimated by taking account of every outcome that would generate a cycle, so it is estimated to be

$$\frac{47.30!}{16.77!\,29.53!}\sum_{i=26}^{50}\left[\frac{50!}{i!\,(50-i)!}\frac{(16.77+i)!\,(79.53-i)!}{97.30!}\right]. \tag{9.17}$$

In the limit as the number of voters approaches infinity, there is a cycle if and only if z is greater than 0.5, so, in view of equation 9.1, the estimate of the probability of a cycle in the limit as the number of voters approaches infinity becomes

$$\frac{47.30!}{16.77!\,29.53!}\int_{0.5}^{1} z^{16.77}(1-z)^{29.53}dz\,. \tag{9.18}$$

With these equations, it is straightforward to compute the estimated probability of a simple tie somewhere in the ranking, the estimated probability the highest-ranked candidate will be either a majority-rule equilibrium (involved in no cycles) or a dominant candidate (involved in neither semi-cycles nor cycles), and the estimated

probability that the entire ranking will be either free of cycles or free of all anomalies (both semi-cycles and cycles). The results of calculations of this sort for selected values of the number of candidates and the number of voters are shown in Tables 9.9 through 9.14.

The Significance of Cycles

There is an important lesson to be drawn from the collection of numbers in Tables 9.9 through 9.14. When there are more than two candidates, there is a significant probability of a cycle that will keep "majority rule" from being well defined. However, that probability is often not much larger than the probability of a tie, which can keep majority rule from being well defined even if there are only two candidates. Thus, for example, when there are two candidates and 150 voters, the estimated probability of a tie is slightly larger than 1 percent. When there are three candidates and more than 1,000 voters, the estimated probability of a cycle is smaller than this. Even when there are 10 candidates and only 200 voters, the 9.9 percent estimated probability that there will be no dominant candidates is about the same as the 9.9 percent estimated probability of a tie when there are two candidates and 20 voters. Thus it can be argued that the possibility of cycles deserves to be put in the same category with the possibility of ties. In the same way that we append devices for breaking ties to majority rule, we can append devices for cutting through cycles to the process of seeking a collective ranking of more than two candidates by majority rule. While cycles will sometimes occur, much more often than not there will be a dominant candidate. It is necessary to be prepared for cycles; it is not necessary to abandon completely the idea of majority rule when there are more than two candidates.

The next chapter interprets Kenneth Arrow's famous theorem as an elaboration of the idea that because cycles can occur, it is never possible to be sure that a collectivity's relative ordering of a given set of options would be preserved if additional options were considered.

Table 9.9 Estimated probability of a tie between the top two candidates

Number	Number of candidates									
of voters	2	3	4	5	6	8	10	15	20	30
4	0.3248	0.3362	0.3371	0.3380	0.3395	0.3435	0.3476	0.3554	0.3604	0.3659
6	0.2536	0.2664	0.2675	0.2684	0.2702	0.2748	0.2795	0.2887	0.2947	0.3013
8	0.2087	0.2222	0.2233	0.2244	0.2262	0.2312	0.2363	0.2465	0.2531	0.2606
10	0.1772	0.1911	0.1923	0.1933	0.1953	0.2004	0.2058	0.2166	0.2237	0.2318
20	0.0990	0.1122	0.1133	0.1143	0.1163	0.1215	0.1271	0.1390	0.1472	0.1571
30	0.0668	0.0782	0.0793	0.0802	0.0819	0.0868	0.0920	0.1035	0.1119	0.1225
40	0.0494	0.0593	0.0602	0.0611	0.0626	0.0670	0.0718	0.0826	0.0908	0.1014
50	0.0387	0.0473	0.0481	0.0489	0.0503	0.0542	0.0585	0.0686	0.0764	0.0869
60	0.0315	0.0392	0.0399	0.0405	0.0418	0.0453	0.0492	0.0585	0.0660	0.0762
70	0.0264	0.0332	0.0339	0.0345	0.0356	0.0387	0.0424	0.0510	0.0580	0.0680
100	0.0175	0.0225	0.0230	0.0235	0.0243	0.0267	0.0296	0.0365	0.0425	0.0514
150	0.0110	0.0144	0.0147	0.0150	0.0156	0.0173	0.0194	0.0246	0.0292	0.0365
200	0.0079	0.0105	0.0107	0.0109	0.0114	0.0127	0.0143	0.0184	0.0221	0.0283
250	0.0061	0.0082	0.0084	0.0086	0.0089	0.0100	0.0113	0.0146	0.0178	0.0231
300	0.0050	0.0067	0.0069	0.0070	0.0073	0.0082	0.0093	0.0121	0.0148	0.0194
350	0.0042	0.0057	0.0058	0.0059	0.0062	0.0070	0.0079	0.0104	0.0127	0.0168
500	0.0029	0.0039	0.0040	0.0041	0.0042	0.0048	0.0054	0.0072	0.0089	0.0119
750	0.0018	0.0025	0.0026	0.0026	0.0028	0.0031	0.0036	0.0047	0.0059	0.0080
1000	0.0014	0.0019	0.0019	0.0020	0.0021	0.0023	0.0026	0.0035	0.0044	0.0060
1500	0.0009	0.0012	0.0013	0.0013	0.0014	0.0015	0.0017	0.0023	0.0029	0.0040
2000	0.0007	0.0009	0.0009	0.0010	0.0010	0.0011	0.0013	0.0017	0.0022	0.0030
2500	0.0005	0.0007	0.0007	0.0008	0.0008	0.0009	0.0010	0.0014	0.0017	0.0024
3000	0.0004	0.0006	0.0006	0.0006	0.0007	0.0008	0.0009	0.0012	0.0014	0.0020
3500	0.0004	0.0005	0.0005	0.0005	0.0006	0.0006	0.0007	0.0010	0.0012	0.0017
limit	0	0	0	0	0	0	0	0	0	0

Table 9.10 Estimated probability of some tie between adjacent candidates

Number	Number of candidates									
of voters	2	3	4	5	6	8	10	15	20	30
4	0.3248	0.5594	0.7169	0.8198	0.8861	0.9550	0.9824	0.9983	0.9998	1.0000
6	0.2536	0.4619	0.6185	0.7323	0.8135	0.9106	0.9575	0.9935	0.9990	1.0000
8	0.2087	0.3951	0.5456	0.6625	0.7511	0.8665	0.9291	0.9857	0.9971	0.9999
10	0.1772	0.3457	0.4890	0.6056	0.6978	0.8251	0.8998	0.9756	0.9941	0.9997
20	0.0990	0.2117	0.3232	0.4260	0.5171	0.6635	0.7684	0.9113	0.9665	0.9953
30	0.0668	0.1503	0.2403	0.3286	0.4115	0.5551	0.6682	0.8452	0.9291	0.9853
40	0.0494	0.1151	0.1900	0.2668	0.3416	0.4775	0.5911	0.7853	0.8896	0.9714
50	0.0387	0.0925	0.1564	0.2240	0.2916	0.4190	0.5301	0.7325	0.8512	0.9550
60	0.0315	0.0768	0.1323	0.1926	0.2541	0.3732	0.4806	0.6861	0.8149	0.9373
70	0.0264	0.0653	0.1144	0.1686	0.2249	0.3364	0.4396	0.6451	0.7809	0.9190
100	0.0175	0.0446	0.0806	0.1220	0.1666	0.2593	0.3501	0.5472	0.6930	0.8638
150	0.0110	0.0286	0.0532	0.0826	0.1155	0.1870	0.2613	0.4371	0.5828	0.7798
200	0.0079	0.0208	0.0394	0.0621	0.0879	0.1459	0.2082	0.3642	0.5029	0.7087
250	0.0061	0.0163	0.0312	0.0496	0.0708	0.1194	0.1729	0.3122	0.4425	0.6489
300	0.0050	0.0134	0.0257	0.0412	0.0592	0.1009	0.1478	0.2733	0.3952	0.5982
350	0.0042	0.0113	0.0219	0.0352	0.0508	0.0874	0.1290	0.2430	0.3571	0.5549
500	0.0029	0.0077	0.0151	0.0244	0.0355	0.0621	0.0932	0.1824	0.2773	0.4561
750	0.0018	0.0050	0.0099	0.0161	0.0236	0.0419	0.0636	0.1288	0.2022	0.3522
1000	0.0014	0.0037	0.0074	0.0120	0.0177	0.0315	0.0483	0.0995	0.1591	0.2870
1500	0.0009	0.0025	0.0049	0.0080	0.0117	0.0211	0.0325	0.0684	0.1116	0.2097
2000	0.0007	0.0018	0.0036	0.0060	0.0088	0.0158	0.0245	0.0521	0.0859	0.1653
2500	0.0005	0.0015	0.0029	0.0048	0.0070	0.0127	0.0197	0.0420	0.0699	0.1365
3000	0.0004	0.0012	0.0024	0.0040	0.0058	0.0106	0.0164	0.0352	0.0589	0.1162
3500	0.0004	0.0010	0.0021	0.0034	0.0050	0.0091	0.0141	0.0303	0.0509	0.1012
limit	0	0	0	0	0	0	0	0	0	0

Table 9.11 Estimated probability of a majority-rule equilibrium

Number	Number of candidates									
of voters	2	3	4	5	6	8	10	15	20	30
3	1	0.7594	0.5858	0.4632	0.3721	0.2461	0.1651	0.0619	0.0232	0.0032
4	1	0.8941	0.8048	0.7338	0.6745	0.5766	0.4960	0.3419	0.2350	0.1099
5	1	0.8046	0.6559	0.5476	0.4641	0.3410	0.2536	0.1221	0.0585	0.0132
6	1	0.9019	0.8183	0.7525	0.6978	0.6072	0.5313	0.3812	0.2720	0.1363
7	1	0.8358	0.7060	0.6098	0.5343	0.4184	0.3309	0.1849	0.1024	0.0307
8	1	0.9105	0.8331	0.7726	0.7224	0.6388	0.5677	0.4222	0.3116	0.1665
9	1	0.8587	0.7435	0.6575	0.5891	0.4814	0.3967	0.2445	0.1490	0.0538
10	1	0.9183	0.8466	0.7908	0.7447	0.6674	0.6006	0.4601	0.3491	0.1967
15	1	0.9011	0.8151	0.7502	0.6978	0.6118	0.5389	0.3896	0.2772	0.1356
20	1	0.9442	0.8920	0.8520	0.8192	0.7632	0.7122	0.5942	0.4889	0.3217
25	1	0.9344	0.8733	0.8271	0.7897	0.7265	0.6695	0.5397	0.4271	0.2578
30	1	0.9574	0.9157	0.8840	0.8582	0.8136	0.7718	0.6696	0.5723	0.4055
35	1	0.9506	0.9024	0.8661	0.8367	0.7863	0.7394	0.6261	0.5205	0.3466
40	1	0.9651	0.9297	0.9030	0.8813	0.8436	0.8076	0.7164	0.6260	0.4634
45	1	0.9600	0.9195	0.8892	0.8647	0.8222	0.7819	0.6809	0.5824	0.4107
50	1	0.9700	0.9388	0.9154	0.8965	0.8633	0.8312	0.7480	0.6631	0.5055
60	1	0.9734	0.9452	0.9241	0.9071	0.8772	0.8479	0.7706	0.6902	0.5374
75	1	0.9730	0.9440	0.9225	0.9052	0.8747	0.8448	0.7653	0.6822	0.5244
100	1	0.9803	0.9585	0.9423	0.9293	0.9062	0.8831	0.8197	0.7508	0.6127
125	1	0.9805	0.9586	0.9424	0.9295	0.9065	0.8834	0.8196	0.7498	0.6095
150	1	0.9838	0.9653	0.9516	0.9407	0.9213	0.9016	0.8462	0.7847	0.6576
175	1	0.9836	0.9647	0.9507	0.9396	0.9199	0.8998	0.8433	0.7803	0.6506
200	1	0.9855	0.9687	0.9564	0.9465	0.9289	0.9109	0.8599	0.8025	0.6823
225	1	0.9852	0.9679	0.9553	0.9452	0.9272	0.9087	0.8565	0.7976	0.6747
250	1	0.9866	0.9708	0.9592	0.9499	0.9335	0.9165	0.8682	0.8135	0.6979
300	1	0.9872	0.9721	0.9611	0.9522	0.9365	0.9203	0.8738	0.8209	0.7086
350	1	0.9877	0.9731	0.9624	0.9539	0.9387	0.9229	0.8778	0.8263	0.7165
425	1	0.9878	0.9732	0.9625	0.9541	0.9389	0.9232	0.8781	0.8266	0.7166
500	1	0.9886	0.9748	0.9648	0.9568	0.9426	0.9278	0.8851	0.8361	0.7310
625	1	0.9887	0.9750	0.9651	0.9572	0.9430	0.9283	0.8859	0.8371	0.7324
750	1	0.9892	0.9762	0.9666	0.9591	0.9456	0.9315	0.8907	0.8438	0.7427
875	1	0.9892	0.9761	0.9666	0.9590	0.9455	0.9314	0.8905	0.8435	0.7421
1000	1	0.9895	0.9768	0.9676	0.9602	0.9471	0.9334	0.8936	0.8477	0.7486
1500	1	0.9898	0.9775	0.9685	0.9614	0.9486	0.9352	0.8964	0.8516	0.7547
2000	1	0.9900	0.9778	0.9690	0.9619	0.9493	0.9361	0.8978	0.8536	0.7577
2500	1	0.9901	0.9780	0.9692	0.9623	0.9497	0.9367	0.8987	0.8547	0.7595
3000	1	0.9902	0.9781	0.9694	0.9625	0.9500	0.9371	0.8993	0.8555	0.7608
3500	1	0.9902	0.9782	0.9695	0.9626	0.9503	0.9373	0.8997	0.8561	0.7616
limt	1	0.9905	0.9788	0.9703	0.9636	0.9515	0.9389	0.9021	0.8595	0.7670

Table 9.12 Estimated probability of a dominant candidate[12]

Number	Number of candidates									
of voters	2	3	4	5	6	8	10	15	20	30
3	1	0.7594	0.5858	0.4632	0.3721	0.2461	0.1651	0.0619	0.0232	0.0032
4	0.6752	0.4147	0.2656	0.1772	0.1213	0.0591	0.0295	0.0054	0.0010	3.E-05
5	1	0.8046	0.6559	0.5476	0.4641	0.3410	0.2536	0.1221	0.0585	0.0132
6	0.7464	0.5189	0.3741	0.2798	0.2141	0.1295	0.0798	0.0243	0.0074	0.0007
7	1	0.8358	0.7060	0.6098	0.5343	0.4184	0.3309	0.1849	0.1024	0.0307
8	0.7913	0.5919	0.4572	0.3650	0.2972	0.2028	0.1404	0.0566	0.0227	0.0036
9	1	0.8587	0.7435	0.6575	0.5891	0.4814	0.3967	0.2445	0.1490	0.0538
10	0.8228	0.6464	0.5223	0.4349	0.3686	0.2713	0.2021	0.0972	0.0463	0.0102
15	1	0.9011	0.8151	0.7502	0.6978	0.6118	0.5389	0.3896	0.2772	0.1356
20	0.9010	0.7915	0.7068	0.6443	0.5939	0.5113	0.4421	0.3046	0.2062	0.0910
25	1	0.9344	0.8733	0.8271	0.7897	0.7265	0.6695	0.5397	0.4271	0.2578
30	0.9332	0.8544	0.7907	0.7432	0.7043	0.6381	0.5789	0.4481	0.3404	0.1889
35	1	0.9506	0.9024	0.8661	0.8367	0.7863	0.7394	0.6261	0.5205	0.3466
40	0.9506	0.8887	0.8372	0.7988	0.7671	0.7121	0.6610	0.5415	0.4353	0.2708
45	1	0.9600	0.9195	0.8892	0.8647	0.8222	0.7819	0.6809	0.5824	0.4107
50	0.9613	0.9100	0.8664	0.8338	0.8069	0.7595	0.7145	0.6050	0.5031	0.3352
60	0.9685	0.9244	0.8861	0.8575	0.8340	0.7921	0.7517	0.6505	0.5532	0.3860
75	1	0.9730	0.9440	0.9225	0.9052	0.8747	0.8448	0.7653	0.6822	0.5244
100	0.9825	0.9529	0.9256	0.9053	0.8886	0.8584	0.8282	0.7482	0.6655	0.5102
125	1	0.9805	0.9586	0.9424	0.9295	0.9065	0.8834	0.8196	0.7498	0.6095
150	0.9890	0.9666	0.9447	0.9285	0.9152	0.8912	0.8667	0.7996	0.7276	0.5855
175	1	0.9836	0.9647	0.9507	0.9396	0.9199	0.8998	0.8433	0.7803	0.6506
200	0.9921	0.9731	0.9539	0.9397	0.9281	0.9071	0.8856	0.8256	0.7599	0.6270
225	1	0.9852	0.9679	0.9553	0.9452	0.9272	0.9087	0.8565	0.7976	0.6747
250	0.9939	0.9769	0.9592	0.9462	0.9357	0.9165	0.8967	0.8412	0.7796	0.6531
300	0.9950	0.9793	0.9627	0.9505	0.9406	0.9226	0.9041	0.8515	0.7928	0.6711
350	0.9958	0.9810	0.9651	0.9535	0.9441	0.9270	0.9093	0.8589	0.8023	0.6841
425	1	0.9878	0.9732	0.9625	0.9541	0.9389	0.9232	0.8781	0.8266	0.7166
500	0.9971	0.9840	0.9694	0.9587	0.9502	0.9346	0.9184	0.8721	0.8195	0.7082
625	1	0.9887	0.9750	0.9651	0.9572	0.9430	0.9283	0.8859	0.8371	0.7324
750	0.9982	0.9863	0.9727	0.9627	0.9548	0.9404	0.9254	0.8822	0.8329	0.7275
875	1	0.9892	0.9761	0.9666	0.9590	0.9455	0.9314	0.8905	0.8435	0.7421
1000	0.9986	0.9873	0.9742	0.9647	0.9570	0.9432	0.9289	0.8873	0.8395	0.7372
1500	0.9991	0.9884	0.9758	0.9666	0.9593	0.9460	0.9323	0.8922	0.8462	0.7470
2000	0.9993	0.9889	0.9766	0.9675	0.9604	0.9474	0.9339	0.8947	0.8495	0.7520
2500	0.9995	0.9892	0.9770	0.9681	0.9610	0.9483	0.9349	0.8962	0.8515	0.7550
3000	0.9996	0.9895	0.9773	0.9685	0.9614	0.9488	0.9356	0.8972	0.8529	0.7570
3500	0.9996	0.9896	0.9775	0.9687	0.9618	0.9492	0.9361	0.8979	0.8538	0.7584
limit	1	0.9905	0.9788	0.9703	0.9636	0.9515	0.9389	0.9021	0.8595	0.7670

Table 9.13 Estimated probability of no cycles anywhere

Number	Number of candidates									
of voters	2	3	4	5	6	8	10	15	20	30
3	1	0.7594	0.4241	0.1708	0.0491	0.0014	1.E-05	8.E-14	1.E-25	2.E-61
4	1	0.8941	0.7006	0.4753	0.2772	0.0589	0.0066	2.E-06	7.E-12	3.E-28
5	1	0.8046	0.5025	0.2382	0.0846	0.0044	0.0001	8.E-12	5.E-22	4.E-53
6	1	0.9019	0.7167	0.4943	0.2933	0.0645	0.0074	2.E-06	8.E-12	2.E-28
7	1	0.8358	0.5622	0.2971	0.1216	0.0092	0.0002	2.E-10	1.E-19	2.E-47
8	1	0.9105	0.7358	0.5190	0.3164	0.0742	0.0092	3.E-06	2.E-11	1.E-27
9	1	0.8587	0.6093	0.3482	0.1580	0.0157	0.0006	2.E-09	8.E-18	2.E-43
10	1	0.9183	0.7541	0.5437	0.3407	0.0856	0.0116	5.E-06	4.E-11	6.E-27
15	1	0.9011	0.7046	0.4653	0.2555	0.0426	0.0031	1.E-07	2.E-14	1.E-35
20	1	0.9442	0.8191	0.6394	0.4437	0.1457	0.0278	4.E-05	1.E-09	2.E-23
25	1	0.9344	0.7886	0.5851	0.3759	0.0967	0.0127	4.E-06	1.E-11	6.E-29
30	1	0.9574	0.8553	0.6981	0.5143	0.1986	0.0471	0.0001	1.E-08	3.E-21
35	1	0.9506	0.8331	0.6562	0.4576	0.1483	0.0268	2.E-05	4.E-10	3.E-25
40	1	0.9651	0.8775	0.7364	0.5636	0.2421	0.0664	0.0003	7.E-08	1.E-19
45	1	0.9600	0.8602	0.7025	0.5152	0.1931	0.0428	0.0001	4.E-09	8.E-23
50	1	0.9700	0.8922	0.7631	0.5996	0.2777	0.0846	0.0006	2.E-07	2.E-18
60	1	0.9734	0.9027	0.7826	0.6268	0.3070	0.1011	0.0009	6.E-07	2.E-17
75	1	0.9730	0.9004	0.7762	0.6154	0.2900	0.0893	0.0006	2.E-07	9.E-19
100	1	0.9803	0.9251	0.8262	0.6910	0.3849	0.1528	0.0029	5.E-06	4.E-15
125	1	0.9805	0.9251	0.8253	0.6884	0.3790	0.1470	0.0025	3.E-06	1.E-15
150	1	0.9838	0.9368	0.8504	0.7286	0.4374	0.1942	0.0058	2.E-05	1.E-13
175	1	0.9836	0.9356	0.8472	0.7229	0.4274	0.1849	0.0049	1.E-05	3.E-14
200	1	0.9855	0.9428	0.8630	0.7490	0.4683	0.2213	0.0085	4.E-05	9.E-13
225	1	0.9852	0.9414	0.8596	0.7429	0.4576	0.2109	0.0073	3.E-05	3.E-13
250	1	0.9866	0.9464	0.8634	0.7403	0.4431	0.1944	0.0054	0.0001	3.E-12
300	1	0.9872	0.9488	0.8760	0.7705	0.5028	0.2540	0.0130	0.0001	9.E-12
350	1	0.9877	0.9505	0.8798	0.7768	0.5134	0.2647	0.0148	0.0001	2.E-11
425	1	0.9878	0.9507	0.8800	0.7770	0.5132	0.2641	0.0146	0.0001	2.E-11
500	1	0.9886	0.9536	0.8866	0.7885	0.5335	0.2855	0.0188	0.0002	7.E-11
625	1	0.9887	0.9539	0.8873	0.7895	0.5351	0.2870	0.0191	0.0002	8.E-11
750	1	0.9892	0.9560	0.8920	0.7977	0.5499	0.3033	0.0228	0.0003	2.E-10
875	1	0.9892	0.9559	0.8918	0.7973	0.5489	0.3021	0.0225	0.0003	2.E-10
1000	1	0.9895	0.9572	0.8947	0.8024	0.5584	0.3128	0.0252	0.0004	5.E-10
1500	1	0.9898	0.9583	0.8974	0.8071	0.5671	0.3227	0.0280	0.0005	9.E-10
2000	1	0.9900	0.9589	0.8988	0.8095	0.5716	0.3279	0.0295	0.0006	1.E-09
2500	1	0.9901	0.9593	0.8996	0.8110	0.5743	0.3310	0.0304	0.0006	2.E-09
3000	1	0.9902	0.9595	0.9001	0.8119	0.5761	0.3331	0.0311	0.0007	2.E-09
3500	1	0.9902	0.9597	0.9005	0.8126	0.5774	0.3346	0.0316	0.0007	2.E-09
limit	1	0.9905	0.9607	0.9028	0.8167	0.5852	0.3440	0.0346	0.0008	3.E-09

Table 9.14 Estimated probability of no anomalies[13] anywhere

Number	Number of candidates									
of voters	2	3	4	5	6	8	10	15	20	30
3	1	0.7594	0.4241	0.1708	0.0491	0.0014	1.E-05	8.E-14	1.E-25	2.E-61
4	1	0.6247	0.2313	0.0491	0.0059	1.E-05	3.E-09	6.E-23	5.E-43	6.E-104
5	1	0.8046	0.5025	0.2382	0.0846	0.0044	0.0001	8.E-12	5.E-22	4.E-53
6	1	0.7073	0.3352	0.1029	0.0201	0.0002	3.E-07	3.E-18	3.E-34	5.E-83
7	1	0.8358	0.5622	0.2971	0.1216	0.0092	0.0002	2.E-10	1.E-19	2.E-47
8	1	0.7611	0.4177	0.1596	0.0416	0.0009	3.E-06	2.E-15	6.E-29	1.E-70
9	1	0.8587	0.6093	0.3482	0.1580	0.0157	0.0006	2.E-09	8.E-18	2.E-43
10	1	0.7991	0.4838	0.2142	0.0680	0.0024	2.E-05	2.E-13	2.E-25	3.E-62
15	1	0.9011	0.7046	0.4653	0.2555	0.0426	0.0031	1.E-07	2.E-14	1.E-35
20	1	0.8914	0.6763	0.4219	0.2121	0.0267	0.0013	7.E-09	7.E-17	5.E-42
25	1	0.9344	0.7886	0.5851	0.3759	0.0967	0.0127	4.E-06	1.E-11	6.E-29
30	1	0.9269	0.7650	0.5446	0.3281	0.0687	0.0066	5.E-07	2.E-13	1.E-33
35	1	0.9506	0.8331	0.6562	0.4576	0.1483	0.0268	2.E-05	4.E-10	3.E-25
40	1	0.9448	0.8141	0.6212	0.4123	0.1140	0.0161	5.E-06	2.E-11	5.E-29
45	1	0.9600	0.8602	0.7025	0.5152	0.1931	0.0428	0.0001	4.E-09	8.E-23
50	1	0.9553	0.8446	0.6724	0.4741	0.1563	0.0283	2.E-05	3.E-10	6.E-26
60	1	0.9621	0.8651	0.7087	0.5205	0.1937	0.0418	0.0001	2.E-09	1.E-23
75	1	0.9730	0.9004	0.7762	0.6154	0.2900	0.0893	0.0006	2.E-07	9.E-19
100	1	0.9749	0.9058	0.7854	0.6271	0.3005	0.0941	0.0007	2.E-07	7.E-19
125	1	0.9805	0.9251	0.8253	0.6884	0.3790	0.1470	0.0025	3.E-06	1.E-15
150	1	0.9807	0.9253	0.8250	0.6872	0.3755	0.1435	0.0022	3.E-06	4.E-16
175	1	0.9836	0.9356	0.8472	0.7229	0.4274	0.1849	0.0049	1.E-05	3.E-14
200	1	0.9834	0.9347	0.8449	0.7186	0.4199	0.1779	0.0042	9.E-06	1.E-14
250	1	0.9852	0.9414	0.8596	0.7429	0.4576	0.2109	0.0073	3.E-05	3.E-13
300	1	0.9849	0.9402	0.8494	0.7170	0.4069	0.1637	0.0031	2.E-05	1.E-13
325	1	0.9859	0.9438	0.8646	0.7509	0.4695	0.2213	0.0083	4.E-05	6.E-13
350	1	0.9866	0.9463	0.8702	0.7602	0.4847	0.2356	0.0101	0.0001	2.E-12
425	1	0.9878	0.9507	0.8800	0.7770	0.5132	0.2641	0.0146	0.0001	2.E-11
500	1	0.9878	0.9508	0.8801	0.7771	0.5131	0.2638	0.0145	0.0001	1.E-11
625	1	0.9887	0.9539	0.8873	0.7895	0.5351	0.2870	0.0191	0.0002	8.E-11
750	1	0.9887	0.9542	0.8878	0.7903	0.5362	0.2882	0.0193	0.0002	8.E-11
875	1	0.9892	0.9559	0.8918	0.7973	0.5489	0.3021	0.0225	0.0003	2.E-10
1000	1	0.9892	0.9558	0.8916	0.7969	0.5481	0.3012	0.0223	0.0003	2.E-10
1500	1	0.9896	0.9575	0.8954	0.8035	0.5603	0.3148	0.0258	0.0004	5.E-10
2000	1	0.9898	0.9583	0.8972	0.8068	0.5664	0.3219	0.0277	0.0005	8.E-10
2500	1	0.9900	0.9588	0.8984	0.8088	0.5701	0.3262	0.0290	0.0006	1.E-09
3000	1	0.9901	0.9591	0.8991	0.8101	0.5726	0.3291	0.0299	0.0006	1.E-09
3500	1	0.9901	0.9593	0.8997	0.8111	0.5744	0.3312	0.0305	0.0006	2.E-09
limit	1	0.9905	0.9607	0.9028	0.8167	0.5852	0.3440	0.0346	0.0008	3.E-09

Notes

1 This was first demonstrated by Duncan Black in "On the Rationale of Group Decision-Making," *Journal of Political Economy* 56 (1948): 23–34.
2 Edgar K. Browning, "A Note on Cyclical Majorities," *Public Choice* 12 (1972): 111.
3 Richard D. McKelvey, "Intransitivities in Multidimensional Voting Models and Some Implications for Agenda Control," *Journal of Economic Theory* 12 (1976): 472–82.
4 Charles Plott, "A Notion of Equilibrium and Its Possibility Under Majority Rule," *American Economic Review* 57 (1967): 787–806.
5 Gordon Tullock, "Why So Much Stability?" *Public Choice* 37 (1981): 189–202.
6 In Joseph Greenberg, "Consistent Majority Rules Over Compact Sets of Alternatives," *Econometrica* 47 (1979): 627–36, it is shown that when there are k options, cycles can be precluded by requiring a majority greater than $(k - 1)/k$ to make a decision between two options.
7 Peter Bernholz, "Logrolling, Arrow Paradox and Cyclical Majorities," *Public Choice* 15 (1973): 87–95.
8 The data and the documentation for them are available from the author or from the Inter-university Consortium for Political and Social Research in Ann Arbor, Michigan.
9 That this is what Condorcet meant is not universally agreed. The difficulty of interpreting Condorcet is reviewed by Duncan Black, *The Theory of Committees and Elections* (Cambridge: Cambridge University Press, 1958), pp. 159–78. Condorcet's intent is decoded to my satisfaction in H. Peyton Young, "Condorcet's Theory of Voting," *Amer. Pol. Sci. Rev.* 82 (1988), pp. 1231–44. This rule was suggested independently by John Kemeny in "Mathematics without Numbers," *Daedalus* 88 (1959): 571–91.
10 See George Terrell, *Mathematical Statistics* (New York, 1999), p. 163.
11 For a fuller discussion of the beta density function see Howard Raifa and Robert Schlaifer, *Applied Statistical Decision Theory* (Cambridge, 1961), pp. 216–18.
12 Numbers less than 0.0001 are represented in scientific notation. Thus 3.E-05 is 0.00003.
13 A tie between two adjacent candidates is a coincidence but not an anomaly. The probability of neither anomalies nor such ties can be computed by multiplying a number in table 9.14 by (1-x), where x is the corresponding number in Table 9.10.

Chapter 10

The Arrow Theorem

Introduction

In the voting procedures that have been discussed to this point, two "possibilities" are compared by majority rule, where each possibility may be a set of options. All of the options in the possibility that loses are eliminated, and if more than one option remains, two further possibilities, formed from uneliminated options, are compared. It was shown that when such procedures are used, the outcome depends on the way in which the options are organized on the agenda and on the strategic and coalition-forming abilities of the voters.

Many voting theorists have prescribed a different kind of voting procedure: Each voter provides a ranking of *all* of the options under consideration. The rankings are then "processed" by a mathematical function, to determine a winning set of options. The winning set may be either a single option, in which case it is chosen as the outcome of the collective decision procedure, or it may be a set containing two or more options, in which case some tie-breaking procedure must be used before there can be a collective decision.

The most famous work that approaches voting in this way is *Social Choice and Individual Values*, by Kenneth Arrow.[1] This work was first published in 1951, with a revised edition published in 1963. In 1980 it was the most widely cited work in the field of public choice.[2] It is the principal work for which Arrow was awarded the Nobel Prize in Economics in 1972. The main conclusion that Arrow reaches is that it is not possible for a voting procedure to have all of the properties that one might think are essential. The theorem in which this is proven is known as the Arrow theorem.

In the years since Arrow's work was published, his ideas have been interpreted repeatedly, and yet widespread misunderstanding of the meaning of the Arrow theorem persists. There are several reasons for this. Partly it is because the limitations of voting are a sensitive issue that people prefer to avoid. Focusing on the limitations of voting threatens to undermine the legitimacy of democracy. Partly it is because the issue is inherently complex and subject to misunderstanding. Partly it is because the way that Arrow stated his theorem is particularly subject to misunderstanding. And partly it is because insufficient attention has been paid to the question, "Of the various conditions of the Arrow theorem that cannot be satisfied simultaneously, which one do actual voting procedures fail to satisfy?"

Actual voting procedures do not satisfy Arrow's Axioms I and II, which he describes as rationality conditions. These Axioms, as explained in more detail later, say that the ordering of any available set of options will always be consistent with an unchanging ordering of all options, as long as individual preferences do not change. Rationality conditions, in one form or another, appear in numerous variations on the Arrow theorem,[3] and actual voting procedures never satisfy these conditions.

The next section is an exegesis of those parts of Arrow's book that are essential to understand the Arrow theorem. The section that follows provides a variation on the Arrow theorem that applies to collective decisions that are made not by an effort to maximize a social welfare function, as in the Arrow theorem, but rather by processing votes from voters. The final section explains how voting can be an attractive way to make collective decisions even though voting procedures do not satisfy rationality conditions.[4]

The Arrow Theorem[5]

On page 2 of *Social Choice and Individual Values*, Arrow describes his effort as one of asking

> … if it is formally possible to construct a procedure for passing from a set of known individual tastes to a pattern of social decision-making, the procedure in question being required to satisfy certain natural conditions.

Arrow uses language here that applies both to voting and to making collective decisions by the maximization of a Bergson–Samuelson social welfare function. There is a formal near-equivalence between maximizing a social welfare function and using a voting procedure to make a collective decision. Recall from Chapter 3 that a social welfare function expresses social welfare as a function of the utilities of individuals, where the utility of each individual is a function of the way goods are allocated. Arrow adopts, after discussion on pages 9–11, the modern assumption in economics that individual utility functions have only ordinal and not cardinal significance. That is, individuals are not able to meaningfully communicate their utility for options as cardinal numbers, but only as rankings. Thus the maximization of a social welfare function, like using a voting procedure to make a collective decision, is a process in which the input is individual rankings of patterns of coordinated actions and the output is the identification of a chosen pattern (or two or more patterns in the event of a tie). Thus the Arrow theorem identifies limitations of both voting procedures and social welfare functions.

Discussing the way in which a community would choose among three options, Arrow says on pages 2 and 3,

> It is expected that choices of this type have to be made repeatedly, but sometimes not all of the three alternatives will be available. In analogy with the usual utility analysis of the individual consumer under conditions of constant wants and variable price–income situations, rational behavior on the part of the community would mean that the community orders the three alternatives according to its collective preferences once for all, and then chooses in any given case that alternative among those actually available which stands highest on this list. A natural way of arriving at the collective preference scale would be to say that one alternative is preferred to another if a majority of the community prefer the first alternative to the second.

Arrow uses the word "alternative" for what is defined as an "option" in this book, and the word "community" for what is defined in this book as a "collectivity." These pairs of words will be treated as perfect synonyms. Arrow also refers to the alternatives at times as "candidates."

Arrow introduces the possibility of a cycle among three options, *A*, *B*, and *C*, and comments on page 3,

> Then a majority prefer *A* to *B* and a majority prefer *B* to *C*. We may therefore say that the community prefers *A* to *B* and *B* to *C*. If the community is to be regarded as behaving rationally, we are forced to say that *A* is preferred to *C*. But in fact a majority of the community prefer *C* to *A*. So the method just outlined of passing from individual to collective tastes [saying that the community prefers one alternative to another if a majority of the individuals in the community prefer the first alternative to the second] fails to satisfy the condition of rationality, as we ordinarily understand it. Can we find other methods of aggregating individual tastes which imply rational behavior on the part of the community and which will be satisfactory in other ways?

Arrow eventually shows that the answer to the question just asked is, "No." Any method of aggregating individual tastes that is "rational" as Arrow defines the term will not be minimally satisfactory in other ways.

Before going more deeply into Arrow's work, it is useful to examine the reasons for this limitation on collective decisions. The crux of the difficulty lies in Arrow's concepts of social rationality and social preference. Stating Arrow's concept of social rationality again (from page 2):

> … rational behavior on the part of the community would mean that the community orders the three alternatives according to its collective preferences once for all, and then chooses in any given case that alternative among those actually available which stands highest on this list.

For a community to do what Arrow asks of it here, it would need a kind of omniscience or infallibility. It would need some way of *knowing*, absolutely, what its "collective preferences, once for all" truly were. If there is any possibility of doing this, it is only when a community is in complete agreement. There is no basis for asserting that "a community wants" what some of its members do not

want. Communities can make collective decisions. That is, they can coordinate their intentions. They can do this by voting, wherein some of the members of the collectivity accede to what they did not wish or did not think best. *But the fact that the community has made a collective decision by voting does not establish that the community "desires" the chosen option.* In making a collective decision, a community does not establish what it desires; it establishes what it intends to do.

If a community has agreed that certain collective decisions shall be made by voting, then when they vote and make such decisions, it is not and cannot be by consulting their true "collective preferences," whatever those might be. At best, the outcome of a voting procedure represents an estimate of what is best for the community. Like any other estimate, an estimate derived by voting is susceptible to error. It is this unavoidable possibility of error in collective decisions, manifest in the possibility of majority-rule cycles, that causes voting procedures not to satisfy Arrow's rationality conditions. Thus if a community has agreed to choose between A and B by majority rule and a majority prefer A, then it is accurate to say that the community has chosen A but not that the community prefers A. This lack of a meaningful concept of community preference prevents communities from satisfying Arrow's description of behaving rationally.

It might be thought that Arrow's result was a consequence of the possibility that voters respond to requests for evaluations of options not with their true evaluations, but rather with strategic ones. On pages 6 and 7 Arrow states that he will ignore the strategic aspects of voting and assume that all stated preferences are true preferences. Thus Arrow is concerned with demonstrating the impossibility of an adequate, rational voting procedure even if no one ever voted strategically. The assumption that voting is never strategic increases the logical strength of his result.

Now it is time to examine Arrow's formal definition of rationality, which is developed in the context of a "preference relation." On page 12 he says,

> The statement "x is preferred or indifferent to y" will be symbolized by $x\ R\ y$. The letter R, by itself, will be the name of the relation and will stand for a knowledge of all pairs such that $x\ R\ y$.

Like numerous other economists, Arrow takes liberties here with English grammar. It would have been more proper for him to have said, "$x\ R\ y$ means that either x is preferred to y or else there is indifference between x and y." But in any case, we know what he means. He then states Axiom I (page 13):

> AXIOM I: For all x and y, either $x\ R\ y$ or $y\ R\ x$.

Arrow gives this axiom the name "connectedness," but "completeness" would be a more customary term today.

In interpreting Axiom I, it is important to distinguish between its meaning as a statement about individual preferences and its meaning as a condition to be applied

to collective rankings of options. As a statement about individual preferences, Axiom I means that individuals are able to compare options. As a condition to be applied to collective rankings, Axiom I embodies an absolutism and infallibility that are in fact unattainable. It says that either the collectivity will regard x as at least as good as y or it will regard y as at least as good as x. If a collectivity were to pronounce upon the relative merit of x and y only once, on the basis of a fixed amount of information about the preferences of the members of the collectivity, then Axiom I could be satisfied. But as Arrow employs Axiom I, the question of whether $x\ R\ y$ or $y\ R\ x$ is true, or both, is not allowed to vary when the collectivity gains information about the views of its members concerning any other option, z, which might be involved in a majority-rule cycle with x and y. It is for this reason that Axiom I is not satisfied by actual voting procedures.

Arrow's Axiom II is one of transitivity:

> AXIOM II: For all x, y, and z, $x\ R\ y$ and $y\ R\ z$ imply $x\ R\ z$.

Again, it is important to distinguish between the meaning of this axiom as a statement about individual preferences and its meaning as a condition on collective rankings. As a statement about individual preferences it is entirely non-controversial, simply reflecting what we understand preferences to be: If an individual regards x as at least as good as y and y as at least as good as z, then that individual regards x as at least as good as z. However, as a condition to be applied to collective rankings, it entails assuming that if x is ranked as at least as good as y when only these two options are available, and y is ranked as at least as good as z when only these two options are available, then x will be ranked as at least as good as z when only these two options are available. This is a condition that actual voting procedures do not satisfy. Because majority-rule cycles can occur, because any collective ranking of options in a cycle is necessarily inconsistent with majority rule for at least one pair, and because we would regard it as natural to abide by majority rule for that pair if those two options were the only ones that voters had ranked, actual voting procedures do not satisfy Arrow's Axiom II.

Arrow has occasion to employ concepts of strict preference and indifference, in addition to the R relation. He uses the symbol P and I to denote strict preference and indifference respectively, and on page 14 he defines these formally by,

> DEFINITION 1: $x\ P\ y$ is defined to mean not $y\ R\ x$.

and

> DEFINITION 2: $x\ I\ y$ means $x\ R\ y$ and $y\ R\ x$.

Arrow proves some simple lemmas about these definitions, including,

> LEMMA 1: (d) If $x\ I\ y$ and $y\ I\ z$, then $x\ I\ z$.

(f) If $x\ P\ y$ and $y\ R\ z$, then $x\ P\ z$.

By the same logic as Arrow uses to prove Lemma 1(f), one can prove

Lemma 1: (f′) If $x\ R\ y$ and $y\ P\ z$, then $x\ P\ z$.

A Digression on the Equivalence of Preference Relations and Rankings and on the Reasonableness of Requiring Reported Preferences to be Transitive

While Arrow's formal presentation is couched in terms of preference relations, he sometimes treats a ranking of the candidates as synonymous with a preference relation (for example, page 27). But Arrow does not demonstrate formally that a ranking is equivalent to a preference relation. Furthermore, actual voting procedures do not require voters to specify their preference relations, though they may require voters to rank the options under consideration. Thus it is important to establish that a transitive preference relation is equivalent to a ranking of the options.

A ranking by a voter is a list of options under consideration in the order of the voter's expressed preferences. A ranking is complete if it includes all options under consideration. For written descriptions of rankings, the convention will be adopted that parentheses within a ranking denote expressed indifference among the options so enclosed. While Arrow assumed that expressed preferences were always true preferences, there is generally no requirement that the preferences expressed in a ranking be true preferences.

Given a complete ranking, the preference relation that corresponds to it is obtained by the rule that for any pair of options (x, y), if x and y are enclosed together in parentheses in the ranking, then $x\ I\ y$. Otherwise, $x\ P\ y$ if and only if x is ahead of y in the ranking. The relation R is then defined by $x\ R\ y$ if and only if $x\ P\ y$ or $x\ I\ y$. The completeness of the ranking implies that R is complete. The transitivity of "is ahead of," in conjunction with the logic of parentheses, implies that R is transitive.

Given a complete and transitive preference relation, the corresponding complete ranking can be identified by the following algorithm: Take any two options. Call them v and w. Place v and w in the order vw, wv, or (vw) according to whether $v\ P\ w$, $w\ P\ v$, or $v\ I\ w$. Now take a third option, x. Put x in the midst of the options that are already ranked, just ahead of the highest-ranked option y such that $x\ P\ y$, if there is such an option. If there is no option y such that $x\ P\ y$, then put x at the end. Enclose x in parentheses with the option just ahead of x (call it z), and any others enclosed with z, if $x\ I\ z$. Take the remaining options, one by one, and treat them in the same manner as x was treated.

Now it must be shown that this algorithm yields a ranking in which every option is preferred to those ranked after it, apart from the indifference among the options in parentheses. Suppose that the rule holds for all options introduced

before some option, y. Let z be the option, if there is one, ranked immediately after y. Let x be the option, if there is one, ranked immediately before y. From the way that the placement of y is prescribed, $y\,P\,z$ and $x\,R\,y$. Consider any option, w, that is ahead of x, with $w\,P\,x$. Since $x\,R\,y$, Arrow's Lemma 1(f) implies $w\,P\,y$. Consider any option, w, that is ahead of x, with $w\,I\,x$. If $x\,I\,y$ then Arrow's Lemma 1(d) implies $w\,I\,y$. If $x\,P\,y$ then Lemma 1(f′) implies $w\,P\,y$. Consider any option, w, that is after z. Since $y\,P\,z$ and $z\,R\,w$, Arrow's Lemma 1(f) implies $y\,P\,w$. Thus the rule that every option is preferred to those ranked after it, apart from indifference among options in parentheses, holds for y and for every option that is introduced. Thus every complete and transitive preference relation is mapped to a unique complete ranking. The correspondence between complete and transitive preference relations and complete rankings means that there is no real difference between a social welfare function that is a function of preference relations and one that is a function of rankings.

While rankings are inherently transitive, preference relations are transitive by assumption. Thus it is relevant to ask whether the requirement of transitivity that is incorporated in the idea of a ranking represents an unfortunate restriction in the preferences that voters are permitted to report.

In discussing the possibility of intransitivity in a voter's preference relation, two types of intransitivity must be distinguished. These may be labeled intransitivity of preference and intransitivity of indifference. An intransitivity of preference occurs if a voter reports preferences of $x\,R\,y$, $y\,R\,z$ and $z\,R\,x$, with at least one of the first two R's being an instance of preference rather than indifference. Upon encountering such a voter, we would say that he does not understand what preference means. Thus the reporting of voters' preferences need not accommodate the possibility of intransitivity of preference.

An intransitivity of indifference occurs if a voter reports preferences of $x\,I\,y$, $y\,I\,z$ and $x\,P\,z$. This form of inconsistency is somewhat easier to understand, because of the possibility of indiscernible differences adding up to a discernible difference. Nevertheless, it is possible to say to such a person, "Even though you are unable to distinguish between x and y and between y and z, the fact that you regard x to be better than z should permit you to deduce that, from your perspective, x is the best of the three options, z is the worst, and y is somewhere in between."[6] Thus the possibility of indiscernible differences adding up to a discernible difference does not provide a rationale for requiring the reporting of voters' preferences to accommodate intransitivity of indifference. It is reasonable to require that preferences be reported in the form of rankings, which precludes intransitivity in reported preferences.

End of Digression and Return to Arrow

On page 15 Arrow introduces the symbol *S* to denote an "environment," or set of available alternatives. He then defines (rational) choice from environment *S*, denoted by *C*(*S*), relative to the order relation *R*:

> DEFINITION 3: *C(S) is the set of alternatives x in S such that, for every y in S, x R y.*

In words, what is chosen is what is at least as good as everything else that is available. This definition allows for the possibility that what is chosen is more than one (equally good) option.

Arrow next explains (page 17) what the options are, over which the preference relation is defined:

> In the present study the objects of choice are social states. The most precise definition of a social state would be a complete description of the amount of each type of commodity in the hands of each individual, the amount of labor to be supplied by each individual, the amount of each productive resource invested in each type of productive activity, and the amount of various types of collective activity, such as municipal services, diplomacy and its continuation by other means, and the erection of statues to famous men. It is assumed that each individual in the community has a definite ordering of all conceivable social states, in terms of their desirability to him. It is not assumed here that an individual's attitude toward different social states is determined exclusively by the commodity bundles which accrue to his lot under each. It is simply assumed that the individual orders all social states by whatever standards he deems relevant.

What Arrow calls a "social state" is related to a "pattern of coordinated intentions" as discussed in Chapter 1 in connection with the definition of a collective decision. There is an important difference, however. A pattern of coordinated intentions is limited in its scope, while a social state describes every aspect of social reality. In addition, calling the options "patterns of coordinated intentions" rather than "social states" emphasizes that these objects of collective decisions cannot be implemented simply by fiat. All individuals whose actions are included in the coordinated plan must will their assent. While it is possible to influence individual wills by rewards and punishments, there are limits to this influence. These limits reduce the range of options that collectivities find it worthwhile to entertain.

In assuming that each individual has a definite ordering of all conceivable social states, Arrow is following the tradition in economics of assuming that individuals have utility functions that apply to all possibilities. However, this assumption is unrealistic. People do not have time to conceive of all conceivable social states. Economists presume that individuals have utility functions that apply to all possibilities because their behavior includes a variety of systematic patterns. But the utility functions have no independent existence; utility functions are born

in the minds of economic analysts in response to the observation of systematic patterns of behavior.

The assumption about individual preferences that Arrow requires for his conclusion with respect to voting procedures is somewhat weaker. What he requires is that for any set of options that may be presented for ranking, individuals will be able to rank them in terms of attractiveness to themselves, and these rankings will be consistent with the rankings that would have been obtained if other options had been presented to them for ranking. The fickleness of voter opinion should lead us to doubt the validity of such an assumption. But again, because it is a negative conclusion that Arrow eventually reaches, this assumption does not count against him. The Arrow theorem implies that even if individuals could provide individual rankings of the options under consideration that were completely consistent with the rankings that they would have provided of options that might have been under consideration, there would still be problems in aggregating these individual rankings into an adequate rational social ranking.

On page 23 Arrow assigns the symbol R_i to denote the preference relation of individual i. Then he states:

> DEFINITION 4: *By a* social welfare function *will be meant a process or rule which, for each set of individual orderings,* $R_1, \cdots, R_n$ *for alternative social states (one ordering for each individual), states a corresponding ordering of the social states, R.*

The next step in Arrow's analysis is to identify a set of conditions that cannot be met simultaneously by any collective decision procedure that satisfies his rationality axioms. The function of these conditions is to disqualify collective decision procedures that satisfy the rationality conditions but are unsatisfactory for other reasons.

One way of making collective decisions rational would be to eliminate the possibility of cycles. As was seen in Chapter 9, this can be done by assigning positions in a one-dimensional continuum to the options under consideration and requiring individuals to report preferences that are single-peaked in that continuum. Devices of this sort, that would assume away the problem of cycles and hence ensure rationality, are excluded by Arrow's Condition 1 on page 24:

> CONDITION 1: *Among all the alternatives there is a set S of three alternatives such that, for any set of individual orderings* $T_1, \cdots, T_n$ *of the alternatives in S, there is an admissible set of individual orderings* $R_1, \cdots, R_n$ *of all the alternatives such that, for each individual i,* $x R_i y$ *if and only if* $x T_i y$ *for x and y in S.*

Condition 1 is couched in terms of the possibility that some rankings would be simply "inadmissible," not permitted. Condition 1 allows some inadmissibility, as long as there is a set of three options that can be ordered in any way by all individuals, within an admissible set of rankings of all of the options.

Another unsatisfactory way in which social rationality conditions might be met would be by having a social welfare function that responded perversely to changes in individual preferences. Arrow's Condition 2, on page 26, forecloses this possibility:

> CONDITION 2: *Let $R_1, \cdots, R_n$ and $R_1', \cdots, R_n'$ be two sets of individual ordering relations, R and R′ the corresponding social orderings, and P and P′ the corresponding social preference relations. Suppose that for each i the two individual ordering relations are connected in the following ways: for x′ and y′ distinct from a given alternative x, $x' R_i' y'$ if and only if $x' R_i y'$; for all y′, $x R_i y'$ implies $x R_i' y'$; for all y′, $x P_i y'$ implies $x P_i' y'$. Then, if x P y, x P′ y.*

Arrow gives Condition 2 the name "positive association." It corresponds to the condition of monotonicity that appears in Chapter 8 in conjunction with characterizations of majority rule. Condition 2 says that if one or more voters raise their individual rankings of one option while all other relative rankings remain the same, then the collective ranking of the option that rose cannot fall with respect to any other option.

Yet another unsatisfactory way of avoiding "irrationality" would be to suppose that decisions depend not only on the preferences of individuals with respect to the options under consideration, but also on their preferences with respect to options that are not under consideration. If this were permitted, then it would be possible to suppose that every decision would be made in the way that it would have been made if it had been based on voters' rankings of all conceivable options. The rationality conditions would then be satisfied. On page 27 Arrow states a condition that precludes this way of satisfying the rationality conditions:

> CONDITION 3: *Let $R_1, \cdots, R_n$ and $R_1', \cdots, R_n'$ be two sets of individual orderings and let $C(S)$ and $C'(S)$ be the corresponding social choice functions. If, for all individuals i and all x and y in a given environment S, $x R_i y$ if and only if $x R_i' y$, then $C'(S)$ and $C(S)$ are the same.*

Arrow gives Condition 3 the name "independence of irrelevant alternatives." In explaining why Condition 3 is reasonable, Arrow says on page 26,

> Suppose that an election is held, with a certain number of candidates in the field, each individual filing his list of preferences, and then one of the candidates dies. Surely the social choice should be made by taking each of the individual's preference lists, blotting out completely the dead candidate's name, and considering only the orderings of the remaining names in going through the procedure of determining the winner.

This is not a convincing argument. Even if a candidate has died between the time the votes are cast and the time they are counted, the information about voters' rankings of the dead candidate could have a bearing on a collectivity's estimate of which available candidate is best. This information does not have to be ignored.

What can be said in favor of a condition like independence of irrelevant alternatives is that no process could possibly require voters to rank all *conceivable* options. Life is too short. If it were feasible to obtain voters' rankings of all conceivable options, it would be sensible to base all collective decisions on voters' rankings of all of these, whether or not they were available, thereby satisfying the rationality conditions. However, since it is inevitable that voters are only going to rank some of the conceivable options, collective decisions must be based on voters' rankings of the subset of conceivable options that they have ranked. Thus, even though Arrow's argument for Condition 3 is unconvincing, something like Condition 3 is reasonable, to deal with the fact that voters cannot rank all conceivable options.[7]

As a variation on a lack of positive association (Condition 2), one can imagine a social welfare function that did not depend at all on individual rankings. This possibility is the subject of Arrow's Definition 5 on page 28:

> DEFINITION 5: *A social welfare function will be said to be* imposed *if, for some pair of distinct alternatives x and y, x R y for any set of individual orderings* $R_1, \cdots, R_n$, *where R is the social ordering corresponding to* $R_1, \cdots, R_n$.

Social welfare functions that are imposed are precluded by Condition 4 on page 29:

> CONDITION 4: The social welfare function is not to be imposed.

If imposed social welfare functions were permitted, the rationality conditions could be satisfied simply by using functions that did not depend on what voters desired.

A variation on the theme of social welfare functions that are not responsive to voters' preferences is the possibility of a dictatorial social welfare function. This is the subject of Arrow's Definition 6 on page 30:

> DEFINITION 6: *A social welfare function is said to be dictatorial if there exists an individual i such that, for all x and y, x* P_i *y implies x P y regardless of the orderings* $R_1, \cdots, R_n$ *of all individuals other than i, where P is the social preference relation corresponding to* $R_1, \cdots, R_n$.

Social welfare functions that are dictatorial are precluded by Condition 5 on page 30:

> CONDITION 5: *The social welfare function is not to be dictatorial.*

If dictatorial social welfare functions were permitted, the rationality conditions could be satisfied by having the social ranking of options conform to one voter's ranking.

Between pages 48 and 59 Arrow develops the following theorem, which he states on page 59:

> THEOREM 2 (General Possibility Theorem): *If there are at least three alternatives which the members of a society are free to order in any way, then every social welfare function satisfying Conditions 2 and 3 and yielding a social ordering satisfying Axioms I and II must be either imposed or dictatorial.*

Note that Arrow mentions Axioms I and II and Conditions 2 and 3 by name, while including Conditions 1, 4 and 5 by references to their content.

It turned out that there is an error in Arrow's attempted proof of this theorem.[8] The theorem as stated is not true, and the premises must be strengthened slightly in one way or another to yield a true theorem. Arrow acknowledges this in the second edition of *Social Choice and Individual Values*, on page 96. To provide a true theorem, Arrow introduces some new conditions. The first, Arrow's Condition *P* (Pareto consistency) is introduced on page 96:

> CONDITION P: *If $x P_i y$ for all i, then $x P y$.*

In words, if everyone ranks x ahead of y, then x will be ranked ahead of y socially. Condition *P* substitutes for the combination of Conditions 2 and 4.

Arrow also introduces a stronger version of Condition 1, on page 96:

> CONDITION 1′: *All logically possible orderings of the alternative social states are admissible.*

Condition 1′ is often called the condition of universal domain. Where Condition 1 required only that there were three options that could be placed in any order, Condition 1′ requires that any order of *all* options be permitted.

On page 97, Arrow states the following theorem:

> THEOREM 2: *Conditions* 1′, 3, *P, and* 5 *are inconsistent.*

For this theorem to be true, it must be interpreted to be a theorem, like Theorem 2 on page 59, about "every social welfare function … yielding a social ordering satisfying Axioms I and II." That is, a more careful statement of the theorem would be that every social welfare function yielding a social ordering satisfying Axioms I and II fails to satisfy at least one of Conditions 1′, 3, *P*, and 5. The proof that Arrow presents on pages 98–100 is a proof of this theorem.

A key concept employed in the proof is that of a *decisive set*, which Arrow defines on page 98 as follows:

> A set of individuals V is decisive for x against y if x is socially chosen when every individual in V prefers x to y and every individual not in V prefers y to x.

There are two major parts to Arrow's proof. The first part of the proof builds a logical connection between two different concepts of dictatorship, showing that if

some individual, I, is by himself decisive for any single ordered pair of options, then he must be a dictator over all options if the other assumed conditions are to hold. The second part of the proof shows that there must be some individual who is decisive for some pair of options if the other assumed conditions are to hold, making that individual a dictator and establishing that the conditions cannot all hold.

The first part of Arrow's proof (pages 98–100) begins with the following definitions:

(1) *$x \overline{D} y$ means that x is socially preferred to y whenever I prefers x to y, regardless of the orderings of other individuals.*

(2) *$x D y$ means that x is socially preferred to y if individual I prefers x to y and all other individuals have the opposite preference.*

Condition 3 permits this discussion of whether x is socially preferred to y, based only on the preferences of individuals with respect to options x and y. Arrow notes that (1) implies (2), and that (2) is equivalent to the statement that I is a decisive set for x against y.

Arrow's proof of the first part of the theorem then proceeds as follows: Consider the case where there are only three options, x, y and z, and suppose that $x D y$. Suppose I ranks the options x, y, z, while all other individuals prefer y to both x and z, with preferences between the last two options variable. Since I prefers x to y and all others prefer y to x, (2) means that $x P y$. Since all individuals prefer y to z, Condition P means that $y P z$. Therefore, by Axiom II (transitivity), $x P z$. And this is true regardless of the rankings of x and z by individuals other than I. It requires only that $x P_I y$. Thus

(3) $$x D y \text{ implies } x \overline{D} z.$$

Now suppose that I ranks the options z, x, y, while all other individuals prefer both z and y to x. By a similar argument, $x P y$ and $z P x$, so that transitivity implies $z P y$. Thus,

(4) $$x D y \text{ implies } z \overline{D} y.$$

Interchanging y and z in (4) yields

(5) $$x D z \text{ implies } y \overline{D} z.$$

Replacing x by y, y by z and z by x in (3) yields

(6) $$y D z \text{ implies } y \overline{D} x.$$

Since $x \overline{D} z$ implies $x D z$, and $y \overline{D} z$ implies $y D z$, (3), (5) and (6) chained together imply

(7) $$x D y \text{ implies } y \overline{D} x.$$

Interchanging x and y in (3), (4) and (7) yields

$$y D x \text{ implies } y \overline{D} z,$$
$$y D x \text{ implies } z \overline{D} x,$$
$$y D x \text{ implies } x \overline{D} y.$$

Each of these can be chained with (7) to yield

(8) $$x D y \text{ implies } y \overline{D} z, z \overline{D} x \text{ and } x \overline{D} y.$$

Summarizing (3), (4), (7) and (8),

(9) *If $x D y$, then $u \overline{D} v$ is true for every ordered pair (u, v) from the three options x, y, and z.*

Thus dictatorial control of a single paired comparison in just one direction, in conjunction with the rationality conditions, implies that I is a dictator for any paired comparison of the three alternatives.

Arrow employs an argument of Blau's to show how the conclusion can be extended to any number of options.[9] Suppose $a D b$, and let x and y be any two options. If x and y are the same as a and b in either order, then add a third option, c, and apply (9) to deduce that $x \overline{D} y$ when u and v are x and y respectively. If just one of x and y is distinct from a and b, then add this option to a and b to form a triple to which (9) can again be applied to yield $x \overline{D} y$. If both x and y are distinct from a and b, then first add x to a and b and deduce from (9) that $a \overline{D} x$ and therefore $a D x$. Then apply (9) to the triple a, x, y to deduce that $x D y$. Thus $a D b$ for some ordered pair has been shown to imply that $x \overline{D} y$ for all ordered pairs (x, y). That is, I is a dictator. Thus if Condition 5 (along with the other conditions used in this part of the proof) is to hold, it cannot be true that $x D_I y$ for any x, y and I.

The second part of the proof is much briefer. Let V be a set of individuals that is decisive, such that no other decisive set is smaller. By Condition P, the set of all individuals is decisive for any pair, so there must be some decisive set. Let the pair of options (x, y) be such that V is decisive for x over y. If Condition 5 is to be satisfied, set V must contain at least two individuals. Divide V into two parts: V_1, which contains only one individual, and V_2, which contains the rest of the individuals in V. Let V_3 be the set of individuals not in V. Consider the case where the individual in V_1 ranks the candidates x, y and z in that order, the individuals in V_2 rank them in the order z, x, y, and the individuals in V_3 rank them in the order y,

z, x. (Note that these rankings correspond to the traditional cycle among three options.)

Since V is decisive for x against y and all voters in V rank x ahead of y, $x\ P\ y$. It cannot happen that $z\ P\ y$, because then V_2 would be decisive for this pair, and V is the smallest decisive set. Therefore $y\ R\ z$. Since $x\ P\ y$, by Lemma 1 (which relies on Axiom II), $x\ P\ z$. But the one person in V_1 is the only person with this ranking of this pair of options, making him a dictator. Thus the conditions are together impossible.

A Variation on the Arrow Theorem for Collective Decisions without Social Orderings

The Arrow theorem shows the impossibility of a minimally satisfactory voting procedure that would generate, from voters' preference relations, a social ordering relation, R, that satisfied Arrow's rationality conditions. The social ordering relation serves as a guide for collective decisions: Under Arrow's assumption of rational choice, what is collectively chosen from any environment, S, is the option or options x from S such that $x\ R\ y$ for all other options y in S. However, it is not necessary to generate a social ordering relation to make collective decisions by voting. Collectivities can use voting procedures that identify the winner or winners of elections without generating social ordering relations. Thus, parallel to Arrow's definition of a social welfare function, one can define a ranking-based vote-processing rule (RBVPR):

> Definition: *By a* ranking-based vote-processing rule *will be meant a process or rule which, for each sequence of individual orderings,* $R_1, \cdots, R_n$ *of alternative social states in any environment S, identifies a subset of S,* $f(S, R_1, \cdots, R_n)$, *as chosen social states*.

To simplify notation, a sequence of individual orderings $R_1, \cdots, R_n$ will be designated by **R** when its components can be understood unambiguously, so that an RBVPR can be written $f(S, \mathbf{R})$.

There is a variation on the Arrow theorem for RBVPRs, with axioms and conditions modified to eliminate any references to social ordering relations. That theorem is proved in this section.

For RBVPRs, the role of Arrow's Axiom I (completeness) is played by the assumption, embedded in the definition of an RBVPR, that an RBVPR can be applied to any environment. Arrow's Axiom II (transitivity) is replaced by an assumption of consistency in choices made from related environments. The required assumption is that if the set of available options shrinks in such a way that at least part of what was chosen originally is still available, then options are chosen

from the reduced set if and only if they were chosen from the original set. This requirement is specified formally in the following axiom:

> Axiom *CC*: *If f is an RBVPR, S is a set of options, S′ is a subset of S,* **R′** *is the sequence of rankings of S′ contained in* **R** *and* $S' \cap f(S, \mathbf{R}) \neq \varnothing$, *then for any x in S′,* $x \in f(S', \mathbf{R'})$ *if and only if* $x \in f(S, \mathbf{R})$.

An RBVPR that satisfies Axiom *CC* will be said to possess choice consistency.

The connection between Axiom *CC* and Axiom II is established by the fact that the combination of Arrow's Axiom II and Arrow's condition of rational choice implies the following version of choice consistency:

> *If S is a set of options, S′ is a subset of S, and* $S' \cap C(S) \neq \varnothing$, *then for any x in S′,* $x \in C(S')$ *if and only if* $x \in C(S)$.

In view of the close connection between Axiom *CC* and Arrow's Axiom II, it should not be surprising that Axiom *CC*, despite its apparent reasonableness, is an axiom that actual democratic voting procedures do not satisfy. Like Axiom II, it precludes the possibility that the change in available information, resulting from a change in the set of available options, would change a voting-based estimate of which option from a given set was best.

Condition 3 is transformed by replacing references to rational choice with references to the outcomes of RBVPRs:

> Condition 3′: *Let* **R** *and* **R′** *be two sequences of individual orderings and let* $f(S, \mathbf{R})$ and $f(S, \mathbf{R'})$ *be the corresponding outcomes under the RBVPR f. If, for all individuals i and all x and y in a given environment S,* $x R_i y$ *if and only if* $x R_i' y$, *then* $f(S, \mathbf{R})$ and $f(S, \mathbf{R'})$ *are the same*.

In the absence of a social ordering relation, a condition parallel to Arrow's Condition *P* can be defined by the rule that if there is an option that is ranked first by all voters, then it is chosen:

> Condition *P′*: *If x is ranked first, tied with no other option, by all voters in the sequence of rankings* **R** *of environment S, then* $f(S, \mathbf{R}) = \{x\}$.

The combination of Arrow's Condition *P* and Arrow's condition of rational choice implies the following version of condition *P′*:

> *If x is ranked first from environment S by all voters, tied with no other option, then* $C(S) = \{x\}$.

Dictatorship in the absence of a social ordering relation is defined to exist when a voter is able to determine that an option will be chosen by ranking it first, irrespective of the rankings of other voters:

> Definition 6′: *An RBVPR f is* dictatorial *if there exists an individual i, such that for any option x in any environment S, the fact that individual i ranks x first in the sequence of rankings* **R** *implies that* $f(S, \mathbf{R}) = \{x\}$, *regardless of the rankings of all individuals other than i.*

RBVPRs that are dictatorial are precluded by Condition 5′:

> Condition 5′: *An RBVPR is not to be dictatorial.*

Definition 6′ and Condition 5′ exactly parallel Arrow's Definition 6 and Condition 5.

The following variation on the Arrow Theorem for RBVPRs can now be proved:

> Theorem 2′: If an RBVPR satisfies Axiom *CC* then it does not satisfy all of conditions 1′, 3′, *P′* and 5′.

Theorem 2′ is proven by developing a "binary voting outcome relation" that parallels Arrow's social ordering relation, and then establishing an isomorphism between the Axioms and Conditions of Arrow's system and properties of a system founded on the binary social choice relation.

One further notational convention will be useful: Let $\mathbf{R}_{xy}$ denote the sequence of rankings of options *x* and *y* embedded in the sequence of rankings **R** of a larger set of options.

> Definition: *The weak order relation R′ is defined* (*relative to the sequence of rankings* **R**) *by the condition that* $x\ R'\ y$ *if and only if* $x \in f(\{x, y\}, \mathbf{R}_{xy})$.

Parallel to the extension of *R* to *P* and *I*, $x\ P'\ y$ means not $y\ R'\ x$ and $x\ I'\ y$ means $x\ R'\ y$ and $y\ R'\ x$.

> Lemma: *Axiom CC implies that R′ is transitive.*
>
> Proof: Suppose $x\ R'\ y$ and $y\ R'\ z$. Then Axiom *CC* implies (based on $x\ R'\ y$) that $y \equiv f(\{x, y, z\}, \mathbf{R})$, and (based on $y\ R'\ z$) that $z \equiv f(\{x, y, z\}, \mathbf{R})$. Therefore, since $f(\{x, y, z\}, \mathbf{R})$ cannot be empty, $f(\{x, y, z\}, \mathbf{R}) = x$. Therefore, from *CC*, $f(\{x, z\}, \mathbf{R}) = x$.

Now two forms of dictatorship for *R′*, parallel to Arrow's two forms of dictatorship, are defined relative to the sequence of rankings **R**:

(1′) $x \overline{D}' y$ *means that* $f(\{x, y\}, \mathbf{R}_{xy}) = \{x\}$, *whenever I prefers x to y, regardless of the orderings of other individuals*;

(2′) $x D' y$ *means that* $f(\{x, y\}, \mathbf{R}_{xy}) = \{x\}$ *if individual I prefers x to y and all other individuals have the opposite preference*.

Now every step in Arrow's proof has a corresponding step in the system pertaining to *R*′, leading first to the conclusion that (2′) implies (1′), and then that conditions 1′, 3′, *P*′, and 5′ cannot all hold if Axiom *CC* holds.

This variation on the Arrow theorem should make it clear that it is not reasonable to expect an RBVPR to possess choice consistency. To do so, an RBVPR would have to fail to satisfy at least one of the other conditions of the theorem, and all of them are more crucial than choice consistency, and all are satisfied by actual RBVPRs.

The Attraction of Voting in Light of the Arrow Theorem

To see that voting can be an attractive way to make collective decisions despite the Arrow theorem and the variation on it offered in previous section, it is only necessary to remember that voting establishes not what a collectivity should do, but only what the collectivity shall do. To agree to coordinate their intentions through voting, it is not necessary that the members of a collectivity believe that voting identifies the available option that is indisputably best; it is only necessary that they believe that the overall pattern of outcomes that will result from their agreeing that a prescribed set of collective decisions will be made by voting will be better (in terms of whatever combination of selfish and altruistic concerns each individual has) than if they do not agree to make such decisions by voting.

In this context, the role of the Arrow theorem and the variation on it is to warn that there is a limit to what can be expected of any voting procedure. If an RBVPR is to satisfy the conditions of universal domain, independence of irrelevant alternatives, Pareto consistency, and non-dictatorship (all of which are essential), then choice consistency cannot be achieved. The outcomes chosen from different sets of available options, on the basis of any given RBVPR and given sequence of rankings of all of the options, will not necessarily be consistent with one another. But this should not be surprising. It should be expected that voting will entail making choices over pairs of options on the basis of majority rule. And when there is a cycle, any way of identifying a winning option will be inconsistent with majority rule for some pair.

There is yet another perspective from which it should not have been expected that an RBVPR would satisfy choice consistency. When there are just two options and preferences are not unanimous, an RBVPR will select one (presumably by majority rule). But that does not imply that the RBVPR has identified the better

option. All that can be said upon the application of an attractive RBVPR is that it seems more likely than not that the chosen option is the better one. If a third option is offered for consideration, and if each voter submits a new ballot on which all three options are now ranked (the first two options being ranked by each voter in the same order as on his original ballot), an outcome in which the option that was originally rejected is chosen cannot be precluded. With the third option present, more information is generated. To assume that this additional information could not possibly reverse the estimate of which of the first two options was the better of the two would be like assuming in a statistical undertaking that no estimate made on the basis of a given body of data will be inconsistent with an estimate based on an augmentation of the initial data with additional data. When the outcomes of RBVPRs are viewed as estimates rather than pronouncements of truth, choice consistency loses its appeal as a condition that RBVPRs must satisfy.

One possible explanation of the intense interest in the Arrow theorem is that this interest reflects a concern for justifying the coercion that democratic societies employ in seeking to motivate their citizens to abide by their rules. People would like a basis for saying, "Now that we have voted we know the right thing to do; even if you disagree, you have an obligation to do what we all now know to be the right thing. We will tolerate no uncooperativeness." Some people may have thought that the virtues of appropriate RBVPRs provided a basis for such an assertion. Thus the Arrow theorem may leave these people in shocked disbelief. A person who believes that voting can identify the indisputably correct course of action and is concerned about the defensibility of his or her beliefs is required by the Arrow theorem to do some serious re-thinking of political philosophy.

A further important message of the Arrow theorem is that whenever majority-rule cycles are possible, RBVPRs are subject to manipulation by the strategic introduction or removal of options from consideration, so as to affect the estimated relative merit of other options.

The next chapter shows that not only do RBVPRs lack choice consistency, but they also cannot be expected to motivate voters in all circumstances to report their rankings of options honestly.

Notes

1 Kenneth Arrow, *Social Choice and Individual Values* (first edition, New York, 1951; second edition, New Haven, 1963).

2 Elizabeth A. Stafford, "Classic Writings in Social Choice, Fiscal Theory and Policy and State and Local Government Finance," Master's Thesis, Economics Department, Virginia Polytechnic Institute and State University (1980), p. 9.

3 For a survey, see Jerry S. Kelly, *Arrow Impossibility Theorems* (New York, 1978) or Kotaro Suzumura, *Rational Choice, Collective Decisions and Social Welfare* (Cambridge, 1983).

4 One early critic of Arrow's assumption of social rationality was James Buchanan, "Social Choice, Democracy, and Free Markets," *Journal of Political Economy* 62 (1954): 114–23.
5 All references in the remainder of this chapter, unless otherwise noted, are to Kenneth Arrow, *Social Choice and Individual Values* (Second Edition, New Haven, 1963). Only page numbers in the text will be given for these.
6 For further discussion of this point, see I.J. Good and T.N. Tideman, "The Relevance of Imaginary Alternatives," *Journal of Statistics and Computer Simulation* 12 (1981): 313–15.
7 For further discussion of this point, see Georges Bordes and Nicolaus Tideman, "Independence of Irrelevant Alternatives in the Theory of Voting," *Theory and Decision* 30 (1991): 163–86.
8 Julian Blau, "The Existence of Social Welfare Functions," *Econometrica* 25 (1957): 302–13.
9 Ibid, p. 310.

Chapter 11

Strategic Voting and the Gibbard–Satterthwaite Theorem

Introduction

One of the most important understandings about RBVPRs (ranking-based vote processing rules) to flow from Arrow's work is the Gibbard–Satterthwaite theorem. This theorem was first conjectured by William Vickrey.[1] It was subsequently proved concurrently by Allan Gibbard and Mark Satterthwaite.[2] While the proof of the theorem is quite complex, the theorem itself can be stated relatively simply as follows:

> The Gibbard–Satterthwaite Theorem: No non-dictatorial RBVPR for more than two options with a universal domain can motivate voters in all circumstances to report their rankings of the options honestly.

An intuitive understanding of why the Gibbard–Satterthwaite theorem should be true can be obtained by considering a typical example of a cycle in some detail. Example 11.1 below shows a cycle:

Example 11.1 A cycle for illustrating the Gibbard–Satterthwaite theorem

2	3	4
x	*y*	*z*
y	*z*	*x*
z	*x*	*y*

The number at the top of each column is the number of voters who rank the candidates in the order below the number. The example has been made asymmetric, so that it would be natural that the outcome would not be a tie. The application of majority rule to the pairs in Example 11.1 yields the results that *x* beats *y* by 6 votes to 3, *y* beats *z* by 5 votes to 4, and *z* beats *x* by 7 votes to 2. Most RBVPRs would say that *z* is the winner from these rankings. However, for what follows it is not important which option wins as long as there is a single definite winner for any sequence of rankings.

Whichever option wins (say z), there is a set of voters whose last choice is selected (those whose rankings are given by the first column). If these voters were to vote strategically and report the same rankings as those whose second choice is selected (those whose rankings are given by the second column) then, since the two groups together form a majority, the first choice of the group that was copied (y, the first choice of those in the second column) would be an option that beat all others in paired comparisons and would be selected by most RBVPRs. Thus the strategizers from column one would have succeeded in obtaining a more favorable outcome for themselves by their strategic voting.

If a group of voters can obtain a more attractive outcome for themselves by all reporting rankings of R' for the options when they all have true rankings of the options of R, then there is some number, N, that is the minimum number of them that must shift their reported rankings from R to R' to accomplish their strategic purpose. Consider the situation when $N - 1$ have made the strategic shift. Call this sequenced of preferences **R**. If **R** represented the true preferences of the voters, then the strategic shift of just one voter from R to R' would achieve an outcome that that voter found more attractive than the outcome that resulted from reporting his true preferences. In such circumstances that voter would have an incentive to vote strategically. Thus it should not be surprising that the Gibbard–Satterthwaite theorem is true.

The formal proof presented here is based on a proof by Green and Laffont.[3] As the source of their proof, they cite a working paper by Schmeidler and Sonnenshein.[4] The proof employs an assumption that RBVPRs never yield ties. This should not be regarded as a restrictive assumption, however, for the following reason: One of the ways of dealing with ties would be to begin the voting process by employing some random device to decide how each possible tie would be resolved if it should arise. Then, having dealt with potential ties in this even-handed way, the voting procedure would continue by asking the voters for their rankings of the options and resolving any ties that arose in the decided fashion. The Gibbard–Satterthwaite theorem as presented here shows that such a procedure cannot motivate voters in all circumstances to report their rankings of the options honestly.

The next section presents a formal statement of the Gibbard–Satterthwaite theorem. The succeeding section presents a proof of the theorem. The final section discusses the significance of the theorem.

Formal Statement of the Theorem

A determinate RBVPR is a function, $f(S, \mathbf{R})$, that maps sets of options combined with sequences of voters' rankings of those options to single chosen options. Voter i's ranking is denoted by R_i, and the number of voters is denoted by N.

A determinate RBVPR is manipulable for the sequence of rankings $(R_1, \cdots, R_N)$ if, for some voter i, whose true ranking is R_i, there is an alternative ranking R_i', such that the outcome that results when he reports R_i' (taking the rankings of the other voters as given) is ranked by i (according to R_i) ahead of the outcome that results when he reports R_i. That is, a determinate RBVPR is manipulable if voter i prefers $f(S, R_1, \cdots, R_i', \cdots, R_N)$ to $f(S, R_1, \cdots, R_i, \cdots, R_N)$.

A determinate RBVPR is defined to be strategy-proof if there is no sequence of rankings at which it is manipulable.

A determinate RBVPR is defined to be dictatorial if there is a voter whose first choice is always picked, irrespective of what others want.

The Gibbard–Satterthwaite Theorem: If a determinate RBVPR for more than two options has a universal domain and is strategy-proof, then it must be dictatorial.

The Proof

Lemma 11.1: For a determinate RBVPR to be strategy-proof, it must be the case that whenever there is a sequence of rankings $(R_1, \cdots, R_N)$ and a subset of possible outcomes, B, with the property that every voter prefers every option in B to every option that is not in B, then $f(S, R_1, \cdots, R_N) \in B$.

Proof of Lemma 11.1

Suppose f is strategy-proof, and suppose $f(S, R_1, \cdots, R_N) = \alpha$, which is not in B. Let β be an element of B, and let $(R_1', \cdots, R_N')$ be such that $f(S, R_1', \cdots, R_N') = \beta$. [We know that there is some such $(R_1', \cdots, R_N')$ because B was defined as a subset of possible outcomes.] Now consider the following sequence of sequences of rankings:

$\mathbf{R}_0 = (R_1', R_2', R_3', \cdots, R_N')$
$\mathbf{R}_1 = (R_1, R_2', R_3', \cdots, R_N')$
$\mathbf{R}_2 = (R_1, R_2, R_3', \cdots, R_N')$
. .
. .
. .
$\mathbf{R}_i = (R_1, R_2, R_3, \cdots, R_i, R_{i+1}', \cdots, R_N')$
. .
. .
. .
$\mathbf{R}_N = (R_1, R_2, R_3, \cdots, R_N)$

Each line of this sequence of sequences differs from the previous line by only one voter's ranking. The sequence of sequences has been constructed in such a way that $f(S, \mathbf{R}_0) = \beta$ and $f(S, \mathbf{R}_N) = \alpha$. Let j be the smallest integer k such that $f(S, \mathbf{R}_k)$ is not an element of B. Voter j is the only voter whose reported ranking is different in $\mathbf{R}_j$ than in $\mathbf{R}_{j-1}$. Voter j prefers $f(S, \mathbf{R}_{j-1})$ (an element of B) to $f(S, \mathbf{R}_j)$ (not an element of B). Thus if it should happen that the voters' true preferences were those given by $\mathbf{R}_j$, then f would be manipulable by j at $\mathbf{R}_j$. Therefore, to satisfy the assumption that f is strategy-proof, it must be the case that $f(S, \mathbf{R}_N)$ is actually an element of B.

Proof of the Gibbard–Satterthwaite Theorem

Given a sequence of rankings $(R_1, \cdots, R_N)$ and any pair of options (α, β), define a modified ranking $R_{\alpha\beta i}$ for each R_i, obtained by putting α and β at the top of $R_{\alpha\beta i}$, in the order in which they occur in R_i, followed by all other options in the order in which they occur in R_i.

Call the determinate RBVPR under consideration f. By Lemma 11.1 we know that $f(S, R_{\alpha\beta 1} \cdots, R_{\alpha\beta N})$ must be either α or β, since α and β are ranked ahead of all other options by all voters.

For every pair of options (γ, δ) and every voter i, construct $R_{\gamma\delta i}$ in the same way. Define an order relation, P (a "social ranking"), by the rule that $\gamma P \delta$ if and only if $f(S, R_{\gamma\delta 1}, \cdots, R_{\gamma\delta N}) = \gamma$.

Because, by assumption, f never produces ties, either $\gamma P \delta$ or $\delta P \gamma$. And $\gamma P \delta$ implies that it is not the case that $\delta P \gamma$. Thus P is a complete and asymmetric relation.

Now take every other possible sequence of rankings $(R_1', \cdots, R_N')$, and repeat the process of defining the $R_{\gamma\delta i}$'s for every pair of options and every voter, so that it is possible to say, for any (γ, δ), whether $\gamma P \delta$ or $\delta P \gamma$ when voters report any sequence of rankings $(R_1', \cdots, R_N')$.

Let F be the name of the function that maps any sequence of voters' rankings of options, such as $(R_1, \cdots, R_N)$, to the set of ordered pairs of options (γ, δ), such that $\gamma P \delta$ when the voters report the sequence of rankings of $(R_1, \cdots, R_N)$. Note that:

(1) F permits all possible sequences of rankings of the options to be reported. That is, F has a universal domain.
(2) If all voters rank γ ahead of δ, then, in view of Lemma 11.1, F yields $\gamma P \delta$. That is, F satisfies the condition of Pareto consistency.
(3) If $(R_1, \cdots, R_N)$ and $(R_1', \cdots, R_N')$ are two sequences of rankings in which the options γ and δ are placed in the same order in R_i and R_i' for all i, then F must yield $\gamma P \delta$ if and only if it yields $\gamma P' \delta$, because otherwise a proof analogous to the proof of Lemma 11.1 would show that f is manipulable. Thus F satisfies independence of irrelevant alternatives.

Now it must be shown that P must be transitive. To do this, suppose that P is not transitive. Then there is some sequence of rankings, $(R_1, \cdots, R_N)$ and some triple of options (α, β, γ), such that

(4a) $\alpha P \beta$, meaning that $\alpha = f(S, R_{\alpha\beta 1} \cdots, R_{\alpha\beta N})$;
(4b) $\beta P \gamma$, meaning that $\beta = f(S, R_{\beta\gamma 1} \cdots, R_{\beta\gamma N})$;
(4c) $\gamma P \alpha$, meaning that $\gamma = f(S, R_{\alpha\gamma 1} \cdots, R_{\alpha\gamma N})$.

Define $(R_1', \cdots, R_N')$ in terms of $(R_1, \cdots, R_N)$ by moving α, β and γ to the top of each ranking while preserving their order in the ranking, and following these by all other options with their order in the ranking preserved. Without loss of generality, assume that $\alpha = f(S, R_1, \cdots, R_N)$.

Define $(R_1'', \cdots, R_N'')$ as the sequence of rankings that is obtained by moving β to the third position in each of the R_i' rankings, and leaving the order of all other options unchanged.

Every ranking R_i places α ahead of γ if and only if the ranking R_i'' places α ahead of γ. Since $\gamma P \alpha$ [by (4c)] and F satisfies independence of irrelevant alternatives, it must be the case that $f(S, R_1'', \cdots, R_N'') = \gamma$.

Consider the following sequence of sequences of rankings:

$\mathbf{R}_0' = (R_1', R_2', R_3', \cdots, R_N')$
$\mathbf{R}_1' = (R_1'', R_2', R_3', \cdots, R_N')$
$\mathbf{R}_2' = (R_1'', R_2'', R_3', \cdots, R_N')$
. .
. .
. .
$\mathbf{R}_i' = (R_1'', R_2'', R_3'', \cdots, R_i'', R_{i+1}', \cdots, R_N')$
. .
. .
. .
$\mathbf{R}_N' = (R_1'', R_2'', R_3'', \cdots, R_N'')$

We know that $f(S, \mathbf{R}_0') = \alpha$, that $f(S, \mathbf{R}_N') = \gamma$, and that $f(S, \mathbf{R}_i')$ is either α or β or γ for all i. Let j be the smallest integer k for which $f(S, \mathbf{R}_k')$ is β or γ. $\mathbf{R}_j'$ and $\mathbf{R}_{j-1}'$ differ only in that R_j'' occurs in $\mathbf{R}_j'$ and R_j' occurs in $\mathbf{R}_{j-1}'$. If $f(S, \mathbf{R}_j') = \gamma$, then f is manipulable at $\mathbf{R}_j'$ if α is ranked ahead of γ in R_j (and therefore in R_j''), and f is manipulable at $\mathbf{R}_{j-1}'$ if γ is ranked ahead of α in R_j (and therefore in $\mathbf{R}_j'$). On the other hand, if $f(S, \mathbf{R}_j') = \beta$, then f is manipulable at $\mathbf{R}_j'$ because β is in third place in R_j'', and substituting R_j results in β being chosen, which is in first or second place in R_j''. Thus the assumption that P was not transitive led to the conclusion that f was manipulable. Therefore if f is to be strategy-proof as assumed, then P must be transitive.

Since *F* satisfies the completeness and transitivity axioms of the Arrow theorem, as well as the theorem's conditions of universal domain, independence of irrelevant alternatives and Pareto consistency, it must, by the Arrow theorem, be a dictatorial social welfare function. The dictator under the social welfare function *F* is a dictator for the determinate RBVPR *f*, which proves the theorem.

Discussion

Since RBVPRs, virtually without exception, satisfy the non-dictatorship and universal domain conditions, a more natural way to express the Gibbard–Satterthwaite theorem is that any non-dictatorial RBVPR for more than two options with a universal domain cannot be strategy-proof. It must, under some configuration of preferences, give some voter an incentive to misstate his preferences.

Remember that the proof of the Gibbard–Satterthwaite theorem, like the proof of the Arrow theorem on which it draws, requires the existence of cycles. As will be brought out in Chapter 13, there are some RBVPRs that offer no incentive for strategic voting in the absence of cycles. If one of these RBVPRs is used, the practical significance of the Gibbard–Satterthwaite theorem depends crucially on the frequency of cycles being more than infinitesimal.

The Arrow theorem and the Gibbard–Satterthwaite theorem identify properties that would be attractive but cannot be expected in RBVPRs. The next chapter discusses attractive properties that can be found in at least some potentially interesting RBVPRs**.**

Notes

1 William Vickrey, "Utility, Strategy, and Social Decision Rules," *Quarterly Journal of Economics* 74 (1960): 507–35.

2 Allan Gibbard, "Manipulation of Voting Schemes: A General Result," *Econometrica* 41 (1973): 587–601; Mark Satterthwaite, "Strategy-proofness and Arrow's conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions," *Journal of Economic Theory* 10 (1975): 187–217.

3 Jerry Green and Jean-Jacques Laffont, *Incentives in Public Decision-Making* (Amsterdam, 1979), pp. 16–19.

4 D. Schmeidler and H. Sonnenshein, "The Possibility of a Cheat-Proof Social Choice Function: A Theorem of A. Gibbard and M. Satterthwaite," Discussion paper No. 89, Economics Department, Northwestern University, 1974.

Chapter 12

Criteria for Evaluating Ranking-Based Vote-Processing Rules

Introduction

Previous chapters introduced a number of properties that might be sought in ranking-based vote-processing rules (RBVPRs). The characterizations of majority rule in Chapter 8 involved the properties of universal domain, anonymity, neutrality, positive responsiveness, monotonicity, and resolvability. In Chapter 10, the variation on Arrow Theorem for RBVPRs introduced the additional properties of choice consistency, independence of irrelevant alternatives, Pareto consistency, and non-dictatorship. In Chapter 11, the Gibbard–Satterthwaite theorem introduced the further property of strategy-proofness. Table 12.1 below sorts these eleven properties into four that are possessed by almost all potentially interesting RBVPRs, one that is not relevant for the evaluation of RBVPRs, two that are possessed by no such RBVPRs, and four that are possessed by some but not other potentially interesting RBVPRs.

Table 12.1 A classification of previously encountered properties of RBVPRs by frequency of occurrence in potentially interesting rules

Almost all rules	Not relevant for evaluating rules	No rules	Some rules
Anonymity	Independence of irrelevant alternatives	Choice consistency	Universal domain
Neutrality		Strategy-proofness	Positive responsiveness
Pareto consistency			Monotonicity
Non-dictatorship			Resolvability

This classification of properties requires explanation.

An RBVPR is anonymous if it does not matter who cast which vote. The outcome depends only on how many rankings of each type were reported. More formally, anonymity in an RBVPR is the property that any permutation of a sequence of rankings yields the same winning set. An RBVPR is neutral if the

conditions that options must meet to be in the winning set are the same for all options. More formally, if an RBVPR is neutral, then interchanging the positions of any two options in all rankings yields an interchange of the winning statuses of the two options. The properties of anonymity and neutrality are generally required for a rule to be perceived as fair. Anonymity is possessed by all RBVPRs that are actually proposed, other than those with deliberately unequal weights for votes. Voting procedures often give a chairman or some other designated person the power to break ties, which is akin to a lack of anonymity in an RBVPR. However, if anonymity is desired in the breaking of ties, a random tie-breaking procedure can be used. Neutrality is not possessed by rules that give preference to the status quo. However, such preference is never an essential part of an attractive RBVPR, but rather an adaptation to circumstances that make a preference for the status quo seem warranted. Thus it is reasonable to confine a search for attractive RBVPRs to rules that possess anonymity and neutrality.

Arrow defined independence of irrelevant alternatives (IIA), Pareto consistency and non-dictatorship as properties of social welfare functions, which yield social rankings of options, as opposed to the winning sets that RBVPRs yield. Another difference between Arrow's framework and that of RBVPRs is that Arrow assumed that there would be no discrepancies between voters' expressed preferences and their true preferences.[1] Thus there may be some uncertainty about how these properties should be defined for RBVPRs.

The most straightforward adaptation of Arrow's condition of Pareto consistency to RBVPRs is that if all voters rank an option first, then that option is chosen. All RBVPRs that are actually proposed possess this property, and no RBVPR that lacked this property could be interesting.

A plausible definition of "dictatorial" for RBVPRs is that an RBVPR is dictatorial if there is a voter whose first choice is always chosen, irrespective of what others want. Some authors have explored the attractiveness of "voting rules" that combine dictatorship with random process. That is, it has been suggested that a random process might be used to decide which voter's views shall prevail.[2] However, such a collective decision procedure should be regarded as a combination of random process and authority, not an example of voting. Without such an element of random process, a dictatorial RBVPR would simply be an unnecessarily complex way of making collective decisions by the mode of authority. It would not be at all interesting in the search for attractive RBVPRs.

For these reasons it will be taken for granted that all of the RBVPRs that come under discussion possess all of the properties in the first column of Table 12.1: anonymity, neutrality, Pareto consistency and non-dictatorship.

One plausible definition of independence of irrelevant alternatives (IIA) for RBVPRs is that an RBVPR possesses IIA if, when voters have ranked options that are not feasible, this information is ignored in applying the RBVPR. As was mentioned in Chapter 10, whether an RBVPR ought to possess this property can be debated, but there is certainly no difficulty in ensuring that RBVPRs do possess it.

There is an alternative interpretation of IIA, that an RBVPR possesses IIA if the winning set is unaffected by changes in voters' preferences with respect to options that voters have not ranked. Every RBVPR possesses this property by virtue of being a function of reported rankings. Since there is no difficulty in ensuring that RBVPRs possess IIA under either interpretation, and it is not clear that one should insist that they possess the version of IIA that they might not possess, IIA is classified as not relevant for the evaluation of RBVPRs and will not be discussed further.

The central message of the Arrow Theorem is that RBVPRs that possess more essential properties will not possess choice consistency. The central message of the Gibbard–Satterthwaite theorem is that RBVPRs that possess more essential properties will not be strategy-proof. Thus it will be taken for granted that the RBVPRs that come under discussion do not possess either of the properties in the third column of Table 12.1, choice consistency and strategy-proofness.

It is properties such as those in the fourth column of Table 12.1, which are possessed by some RBVPRs and not others, that have the potential for distinguishing between more attractive and less attractive RBVPRs. But the list in the third column of Table 12.1 contains only properties that happened to arise in Chapter 8, in discussions of May's characterization and my characterization of majority rule. The full list of properties that will be used to evaluate RBVPRs is shown in Table 12.2 below:

Table 12.2 A taxonomy of properties used to evaluate RBVPRs

Domain	Consistency	Responsiveness	Stability	Qualitative attractiveness
Universal domain	Majority consistency	Positive responsiveness	Monotonicity	Lucidity
	Mutual majority consistency	Resolvability	Homogeneity	Ease of use
	Condorcet consistency	Perfect-tie responsiveness	Combinativity	Cost of computation
	Smith consistency		Independence of clones	Resistance to strategy
	Schwartz consistency			
	Invariant loss consistency			

The property of universal domain concerns the rankings that voters are permitted to report and for which a rule yields a result. The properties in the

consistency category are concerned with consistency between the processing of rankings by an RBVPR and the processing of pairs of options by majority rule. The properties in the responsiveness category are concerned with the way that an RBVPR responds to small changes in one voter's ranking of one option or to the addition of a single additional ranking, when it is arguable that such a change should resolve a tie in favor of a particular winner. The properties in the stability category are concerned with ensuring that the winning set is not changed inappropriately by changes in the sequence of rankings.

The properties in the previous four categories are defined in such a way that an RBVPR either possesses them or does not possess them. They are not possessed in varying degrees, so there should be no dispute about whether an RBVPR possesses them, once a valid proof had been offered. The properties in the "qualitative attractiveness" category, by contrast, are possessed in degrees and are susceptible to disagreement about the extent to which RBVPRs possess them. The only other feature that unites these properties is that they are desired in RBVPRs. The five sections that follow discuss the properties in the five columns of Table 12.2. The last section summarizes the chapter.

Domain

> *Definition*: Universal domain in an RBVPR is the property that the RBVPR yields a result (though possibly a tie) for every possible sequence of rankings of the options on any agenda.

There are at least four distinguishable ways in which an RBVPR can fail to possess universal domain.

First, there might be some rankings of the options that voters were not permitted to report. This is not as far-fetched as it might at first seem. Recall the discussion of single-peakedness in Chapter 9. If there is reason to believe that voters' rankings of options are invariably single-peaked for a known ordering of the options, then requiring voters to report rankings that are single-peaked for this ordering of the options precludes only strategic votes while it eliminates the possibility of cycles. Thus it is possible that an RBVPR that lacked universal domain would be attractive in some circumstances. However, since in most cases there is no ordering of the options for which all voters' rankings can be expected to be single-peaked, a lack of universal domain cannot generally be justified in this way.

A second way in which an RBVPR can fail to possess universal domain is that voters might not be permitted to report ties in their rankings. This might be done to simplify the processing of rankings, even though the rule in use could accommodate ties if they were permitted.

A third way in which an RBVPR can fail to possess universal domain is that voters might be required to report ties in their rankings. When the customary plurality rule is interpreted as an RBVPR, it is a rule under which voters must report all options except one as tied for second place. One can imagine an alternative version of the plurality rule under which voters report full rankings and the vote-counting process uses only the first-place information. In the next chapter, the plurality rule will be analyzed under both interpretations.

The fourth way in which an RBVPR might fail to possess universal domain is that it might permit voters to report all rankings but leave some combinations of rankings outside the domain of the rule. For example, one might exclude from the domain all combinations of rankings that entailed cycles. This would amount to ignoring the real phenomenon of cycles. Any RBVPR that lacked universal domain for this reason, entailing a possibility of not being able to specify a winning set of options for some combinations of rankings that were individually permitted, would not be at all interesting.

Consistency

The properties in the consistency category pertain to consistency between the outcomes of an RBVPR and majority rule.

Majority consistency is the property that an option that is ranked first by more than half of the voters must win the election. More formally,

> *Definition*: Majority consistency in an RBVPR, f, is the property that whenever there is an option, x, that is ranked first in more than half of the rankings in some sequence of rankings $\mathbf{R}$ for some environment S, $f(S, \mathbf{R}) = \{x\}$.

Mutual majority consistency is the property that if more than half of the voters rank all of the elements of some subset of the options, T, ahead of all other options, all of the winners of the election must be elements of T. More formally,

> *Definition*: Mutual majority consistency in an RBVPR, f, is the property that whenever there is a subset, T, of some environment, S, such that more than half of the rankings in some sequence of rankings, $\mathbf{R}$, place all of the elements of T ahead of all other elements of S, $f(S, \mathbf{R}) \subset T$.

Note that mutual majority consistency implies majority consistency.

The next three properties in the consistency category are related to the idea of dominance.

> *Definition*: An option is dominant if and only if it beats every other option under consideration in head-to-head comparisons by majority rule.

Definition: Condorcet consistency in an RBVPR, f, is the property that whenever some option, x, is dominant for some sequence of rankings, **R**, in some environment, S, $f(S, \mathbf{R}) = \{x\}$.

Note that Condorcet consistency implies majority consistency.

The property of Condorcet consistency is named for the Marquis de Condorcet, who first proposed it in 1785.[3] A dominant option is often called a "Condorcet winner," but that term will not be used here, to avoid confusion with "the winner under the Condorcet rule" (to be defined in Chapter 13).

One of the reasons that Condorcet consistency is a valuable property in RBVPRs is that it greatly reduces the scope for strategic voting. If voters believe that there is virtually no chance that cycles will occur, then the only way in which their votes might affect the outcome is by determining which of two options will be dominant. If voters have such expectations and the RBVPR in use is Condorcet consistent, then they have no incentive to vote strategically. When there is no dominant option, one still might expect an RBVPR to select as the winner one of the options from the "top cycle." John Smith defined the Condorcet criterion as the rule that when every options in set T has a majority over every option in set U, the social ranking should place every option in set T ahead of every options in set U.[4] This implies that the winning option should come from the smallest set such that everything in the set beats everything outside the set by majority rule.[5] This leads to the following definition:

Definition: The Smith set, T, for the sequence of rankings **R** in environment S is the smallest subset of S such that every option in T beats every option not in T when pairs of options in S are compared by majority rule on the basis of the rankings **R**.

The idea of the Smith set leads to the idea of Smith consistency:

Definition: Smith consistency in an RBVPR, f, is the property that for any sequence of rankings, **R**, in any environment, S, every element of $f(S, \mathbf{R})$ is an element of the Smith set.

The following algorithm can be used to identify the Smith set: Identify the option or options that beat the largest number of other options in head-to-head comparisons by majority rule, add all options that beat or tie these, all that beat or tie these, and so on until no more are added. This set is the Smith set.

Smith consistency implies Condorcet consistency since a dominant option, when there is one, is the sole element of the Smith set.

Schwartz proposed that the winning outcome ought to come from a set that can be smaller than the Smith set when there are ties. He called this set the "GOCHA set" (Generalized Optimal-CHoice Axiom).[6] The definition of the GOCHA set

requires some preliminary definitions. Define a set of options, G, to be undominated for the rankings **R** in environment S if, according to **R**, no option in S outside G beats any option in G in head-to-head comparisons by majority rule. Define a set of options to be minimal undominated if it is undominated and it has no proper subset that is undominated.

Definition: The GOCHA set for any sequence of rankings, **R**, in any environment, S, is the union of minimal undominated sets of options.

The idea of the GOCHA set leads to the idea of Schwartz consistency:

Definition: Schwartz consistency in an RBVPR, f, is the property that for any sequence of rankings, **R**, in any environment, S, every element of $f(S, \mathbf{R})$ is an element of the GOCHA set.

The following algorithm can be used to identify the GOCHA set: For any option x, define $W(x)$ (the generalized winners over x) as the union of x with all options that beat x and all that beat them, and so on until no more are added. Define $L(x)$ (the generalized losers to x) as the union of x with all options that lose to x and all that lose to them, and so on until no more are added. Option x is in the GOCHA set if and only if $W(x) \subseteq L(x)$.

To see why this algorithm works, another definition is needed. Define $E(x)$ as $W(x) \cap L(x)$. Each $E(x)$ is a set of options in a cycle or a single option not in any cycle. The condition that $W(x)$ be a subset of or equal to $L(x)$ is the condition that there be no elements of $W(x)$ that are not in $E(x)$. That is, the GOCHA set consists of those cycles or options that are not beaten by outsiders.

Since the GOCHA set is a subset of the Smith set, any RBVPR that is Schwartz consistent is Smith consistent.

The final property in the consistency category, invariant loss consistency, is the property that an option that is beaten by all other options cannot be a winning option.[7] More formally,

Definition: The property of invariant loss consistency in an RBVPR, f, is the property that if option x is beaten by every other option in an environment S in head-to-head comparisons by majority rule when compared in terms of the rankings **R**, then x is not an element of $f(S, \mathbf{R})$.

Responsiveness

The properties in the responsiveness category are concerned with ensuring that ties be susceptible to resolution by appropriate changes in rankings. In Chapter 8, two of these properties were defined with respect to vote processing rules for pairs of

options. Extending the definitions of these properties to RBVPRs is straightforward.

Positive responsiveness is the property that if the performance of an option that is a winner or tied winner improves incrementally, it is then the sole winning option. To specify a formal definition of positive responsiveness, a formalization of the idea of an option moving up incrementally in a ranking is needed:

Definition: The sequence of rankings $\mathbf{R}'$ represents an incremental upward movement for option x relative to the sequence $\mathbf{R}$ if there is another option y and a voter i such that either $x\ P_i'\ y$ and $y\ R_i\ x$ or $x\ R_i'\ y$ and $y\ P_i\ x$, while for all other pairs $\{w, z\}$, $w\ P_i'\ z$ iff $w\ P_i\ z$, and for all voters j other than i, $w\ P_j'\ z$ iff $w\ P_j\ z$, for all pairs $\{w, z\}$.

Definition: Positive responsiveness in an RBVPR, f, is the property that if $x \in f(S, \mathbf{R})$ and the sequence of rankings $\mathbf{R}'$ represents an incremental upward movement for option x relative to the sequence $\mathbf{R}$, then $f(S, \mathbf{R}') = \{x\}$.

Positive responsiveness is an extremely strong property. It says that if there is a tie, then any upward movement of one of the tied options in any ranking, even from last to tied for last, is sufficient to make the option that rose the sole winner of the election. There are few RBVPRs that possess positive responsiveness.

Resolvability is the condition that, for any option involved in a tie, there is some ranking that, when appended to the rankings that created the tie, breaks the tie in favor of that option.

Definition: Resolvability in an RBVPR, f, is the property that if f is an RBVPR and x is an element of $f(S, \mathbf{R})$, then there exists a ranking of the options, Q, such that if $\mathbf{R}'$ is formed by appending Q to $\mathbf{R}$, then $f(S, \mathbf{R}') = \{x\}$.

While it is sometimes difficult to specify how the addition of a particular ranking ought to affect a tie, there are some ties that are so symmetric that it seems possible to specify how the addition of some rankings ought to affect them. Thus it is interesting to define a perfect tie, a tie such as that in Example 9.1, in which the performance of any option is indistinguishable from that of any other option:

Definition: A sequence of rankings is a perfect tie if and only if, for every ordered pair of options (x, y), there is a permutation of the options, π, that maps x to y, and a permutation of the rankings, ρ, that together yield a sequence of rankings that is the same as the original sequence.

In other words, when a sequence of rankings is a perfect tie, it is possible to erase the name of any option x and substitute the name of any other option y on every ranking where it appears, and if one also erases every other name k and replaces it

with $\pi(k)$, and then permutes the sequence of rankings properly, the original sequence will be restored. Under any RBVPR that is anonymous and neutral, all options will be tied winners for a sequence of rankings that is a perfect tie. Thus if such a sequence is augmented by one additional ranking, it is natural to suppose that the option or options at the top of that one additional ranking should be chosen. This is the idea behind of perfect-tie responsiveness:

> *Definition*: Perfect-tie responsiveness in an RBVPR, *f*, is the property that if any sequence of rankings that is a perfect tie is augmented by one additional ranking, the set of winners is the set of options that are ranked first in the one additional ranking.

Perfect-tie responsiveness turns out to be a very strong condition, which few RBVPRs possess.

Stability

The properties in the stability category are concerned with ensuring that the winning set is not changed by changes in the sequence of rankings that arguably should not change it.

Monotonicity is the condition that an incremental upward movement for an option not affect its winning status adversely.

> *Definition*: Monotonicity in an RBVPR, *f*, is the property that if $x \in f(S, \mathbf{R})$, and the sequence of rankings $\mathbf{R}'$ is an incremental upward movement for option x relative to the sequence $\mathbf{R}$, then $x \in f(S, \mathbf{R}')$.

Since monotonicity is a weakening of the condition of positive responsiveness, positive responsiveness implies monotonicity.

A mathematical function of several variables is said to be "homogeneous of degree 0" if multiplying all of the arguments of the function by any constant leaves the value of the function unchanged. By extension, Peyton Young has proposed a definition of homogeneity in voting rules:[8]

> *Definition*: Homogeneity in an RBVPR is the property that the winning set for a given sequence of voters' rankings is the same as the winning set for the sequence obtained by repeating the given sequence any finite number of times.

The attraction of homogeneity should be clear. Nearly all RBVPRs possess homogeneity. While there is at least one RBVPR that has been seriously entertained that does not possess homogeneity (the Dodgson rule), there is a slight

modification of this rule that preserves its essential spirit while also possessing homogeneity.

Combinativity is the property that if two sequences of voters' rankings yield the same winning set, then the sequence of rankings obtained by appending the second to the first should yield that same winning set:[9]

> *Definition*: Combinativity in an RBVPR, f, is the property that if $\mathbf{R}$ and $\mathbf{R}'$ are two sequences of rankings and $\mathbf{R}''$ is the sequence obtained by appending $\mathbf{R}'$ to $\mathbf{R}$, then $f(S, \mathbf{R}) = s$ and $f(S, \mathbf{R}') = s$ implies $f(S, \mathbf{R}'') = s$.

Note that combinativity implies homogeneity.

Combinativity is closely connected to the concept of convexity. The domain of an RBVPR that is anonymous and homogeneous may be conceived as the "election outcome space" of vectors whose components are the fractions of rankings of each possible type. An RBVPR divides election outcome space into regions won by the different options. The region of election outcome space in which each option wins is a convex set if and only if the RBVPR possesses combinativity.

While combinativity may seem to be just a slight strengthening of the property of homogeneity and eminently reasonable at first blush, it turns out to be a very strong property that is inconsistent with the combination of perfect-tie responsiveness and Condorcet consistency.

The proof that no RBVPR can possess all three of the properties of Condorcet consistency, perfect-tie responsiveness and combinativity is as follows: Consider the sequences of ballots in Example 12.1. (A number at the top of a column is the number of voters who rank the candidates in the order below the number.)

Example 12.1

Sequence $\mathbf{R}$:

5	4	4
x	y	z
y	z	x
z	x	y

Sequence $\mathbf{R}'$:

1	4	2	1
x	x	z	z
y	z	x	y
z	y	y	x

Suppose that f is an RBVPR that possesses Condorcet consistency, perfect-tie responsiveness and combinativity, and define $\mathbf{R}''$ to be the sequence obtained by appending $\mathbf{R}'$ to $\mathbf{R}$.

Sequence $\mathbf{R}$, less one ballot from the first column, forms a perfect tie, so the perfect-tie responsiveness of f implies that $f(S, \mathbf{R}) = \{x\}$. For $\mathbf{R}'$, option x is dominant, so the Condorcet consistency of f implies that $f(S, \mathbf{R}') = \{x\}$. Therefore the combinativity of f implies that $f(S, \mathbf{R}'') = \{x\}$. But z is dominant for $\mathbf{R}''$, implying that $f(S, \mathbf{R}'')$ must be $\{z\}$ and establishing a contradiction. Thus no

RBVPR can possess all three of the properties of Condorcet consistency, perfect-tie responsiveness and combinativity. Most voting rules that have been proposed lack both combinativity and perfect-tie responsiveness, but some lack Condorcet consistency, and some lack all three properties.

One way of explaining why combinativity should not be expected is to recall that an RBVPR does not identify infallibly what ought to be done. Rather, it provides a fallible estimate of the appropriate course of action, based on limited information. The information that what had been considered two separate sequences of ballots are actually two parts of the same election is additional information, which can in some cases produce a different estimate of the most appropriate course of action than that produced by considering each sequence separately.

The property of independence of clones requires an elaborate development, beginning with a definition of clones:

> *Definition*: When there are more than two options, a subset C of the options, consisting of at least two options but less than the whole set, is a set of clones (for a given sequence of voters' rankings of the options, **R**) if for any option x that is not in C, every voter either ranks every option in C ahead of x, or ranks every option in C behind x, or ranks every option in C tied with x.

From its usage in biology to denote genetically identical organisms, the word "clones" is intended to suggest that there are no meaningful differences among options that are members of a set of clones. The connection between the idea of virtually identical options and the definition of options that are clones arises from a spatial model of voting.

In the spatial model of voting, each option is characterized by the quantities it possesses of various "attributes," which can be conceived as spatial dimensions. Every voter has an "ideal point" in attribute space, a point that he would most like an option to occupy. Around each voter's ideal point are contours that represent locations of equal attractiveness. A voter ranks all options inside any given contour ahead of all options on or outside the contour. For any pair of options x and y, the set of locations for options that a given voter would rank between x and y would be the region between the contour on which x lay and the contour on which y lay. If x and y have exactly the same attributes (if x and y are indistinguishable, i.e., clones), then, no matter what a voter's preferences in attribute space are, the region of options between x and y consists only of the contour on which x and y lie. It has no width. Thus the probability of finding an option between x and y is less than any positive number.

It could be argued that if two options are clones, then all voters should rank them as tied. But both because some voting rules do not permit voters to rank options as tied and because it is possible that voters would provide complete rankings of options even if they regarded options as equivalent, it is useful to have

a definition of clones that does not require that voters rank clones as tied. Thus the definition of clones employs the weaker condition that no non-clone be ranked ahead of one clone and behind another.

When there is a set of options, C, that satisfies the definition of clones, one cannot be completely sure that the options in C have exactly the same attributes. It could be just a matter of happenstance that other options are never ranked ahead of some elements of C and behind others. However, one can say that such rankings are consistent with the hypothesis that the options in C are arbitrarily close to each other in attribute space. This is the basis of the definition of clones.

The idea behind the concept of independence of clones is that the addition or deletion of clones from the set of options under consideration should have no effect on either the winning status of non-clones or whether clones are among the winners:

> *Definition*: Let S be a set of options. Let C be a subset of S that is a set of clones for the sequence of rankings $\mathbf{R}$. Let x be an element of C. Let $\mathbf{R}'$ be the sequence of rankings obtained by deleting x from every ranking in $\mathbf{R}$. Let y be an option that is not an element of C. An RBVPR f is independent of clones if and only if:
>
> 1. $y \in f(S, \mathbf{R})$ iff $y \in f(S, \mathbf{R}')$, and
> 2. $C \cap f(S, \mathbf{R}) \neq \varnothing$ iff $C \cap f(S, \mathbf{R}') \neq \varnothing$.

The reason that independence of clones in RBVPRs is desirable is related to the spatial model of voting. For any given set of spatial locations of options that voters rank, one would expect an attractive RBVPR not to allow the locations of the selected option to be affected by the number of options from each location that were being ranked. This is what the property of independence of clones prescribes.

While this chapter is concerned primarily with desirable properties that RBVPRs might have, it is worth noting here that the concept of independence of clones has an important analog for tie-breaking procedures. One might reasonably desire a tie-breaking procedure to have the property that the probability of selecting an option from a set of clones would be unaffected by the number of clones in the winning set. This can be achieved by using a random process to select a voter whose ranking will be used to break any ties. If the first voter ranks some options as tied, a second randomly selected voter can be used to break those, and so on until the only remaining ties are those among options that all voters rank as tied, among whom a random process can be used to break ties without violating independence of clones.

Qualitative Attractiveness

The four properties in the qualitative attractiveness category, lucidity, ease of use, cost of computation, and resistance to strategy, are desired features that can be possessed in degrees, as opposed to the properties in other categories, which are either possessed or not possessed. The properties in this category are also, to at least some extent, subjective. Thus it may not be possible to resolve disagreements about which of two RBVPRs possesses them to a greater degree.

Lucidity is a relevant criterion for evaluating RBVPRs because people may be unhappy with a rule that they do not understand, even if experts are convinced of its technical merit. While lucidity is determined to some extent by a society's history, it is determined primarily by the complexity of a rule. Although lucidity is desirable, a low degree of lucidity will not necessarily cause people to regard an RBVPR as unacceptable. Just as people accept devices that they do not understand in the cars and appliances they buy, they may accept an RBVPR they do not understand on the basis of recommendations of experts who claim to understand it.

Ease of use is measured by the time and effort required of voters to provide the information that a voting rule requires of them. All RBVPRs have the same ease of use because they all require voters to submit rankings. The criterion is relevant nevertheless because RBVPRs must be compared with rules such as the plurality rule that require responses from voters other than rankings. There may also be some instances when voters are permitted to submit truncated rankings and others when they are required to submit full rankings, which require more effort.

While the cost of computation of an RBVPR has some relation to its lucidity, it is in principle a different property. It is possible for an RBVPR to be lucid without being computable in a reasonable amount of time, even with modern computers. Suppose, for example, that an RBVPR began with the instruction, "Form the set of all subsets of voters, and for each subset, undertake the following procedure: …" No matter what came next in the rule, modern computers would require an exceedingly long time to determine the result if there were more than a few hundred voters. Sometimes the use of a rule will be feasible if the number of voters or the number of options is small, but not if one or the other is large. Sometimes the use of a rule will be feasible if a computer is available, and not if one is not available. Thus the cost of computation will sometimes be a deterrent to using a rule, and at others times will effectively bar the use of a rule.

No attractive RBVPR will be completely resistant to strategy, since the Gibbard–Satterthwaite theorem shows that all minimally attractive RBVPRs for more than two options will provide incentives for strategic voting in at least some cases. But RBVPRs can vary greatly in the proportion of circumstances in which they offer such incentives and in the strength of the incentives when they do exist. Thus the degree to which an RBVPR resists strategizing is relevant.

In evaluating RBVPRs for resistance to strategy, it will be convenient to note three different types of strategy, which will be described as "burying," "direct

hoisting," and "indirect hoisting."[10] To "bury" an option is to rank it lower than one's truthful ranking, to reduce the probability that the buried option will be elected. To hoist an option is to rank it higher than one's truthful ranking. In the direct hoisting strategy, a voter hoists an option that is regarded as having a significant prospect of winning, even though it is not the voter's first choice, to increase the probability that the hoisted option will defeat an option that is popular with other voters. In the indirect hoisting strategy, a voter hoists an option that is expected to lose, after it eliminates an option that threatens the voter's first choice.

To provide a score for an election rule's resistance to strategy, the following procedure will be used. For each of the 87 elections that were examined in Chapter 9, a computation will be made of the minimum number of voters who would need to cooperate in a strategy to achieve its success. A computation will also be made of the number of voters whose interests would be advanced by the success of the strategy. If the second number is less than the first, the election will be regarded as not susceptible to strategy, and the rule will be given a score of 10 for that election. If the second number is greater than or equal to the first, then a statistical calculation will be made of the expected fraction of the voters that a person seeking to organize a strategy would need to talk to, to find the minimum number needed to make the strategy work. Multiplying this fraction by 10 gives the rule a score for that election in the range of 0 to 10. The rule's overall score for resistance to strategy is the average of the scores for the 87 elections.

Summary

This chapter has defined, discussed and categorized criteria for evaluating vote-processing rules. Criteria for comparing RBVPRs were divided into five categories, concerned with domain, consistency, responsiveness, stability and qualitative attractiveness.

The one criterion in the domain category is universal domain. A rule with universal domain allows voters to report any actual preferences and generates a result for every possible combination of preferences.

The criteria in the consistency category are majority consistency, mutual majority consistency, Condorcet consistency, Smith consistency, Schwartz consistency, and invariant loss consistency. These criteria are all concerned with maintaining such consistency as can be maintained between the outcomes of an RBVPR and the results of comparing pairs of options by majority rule.

The criteria in the responsiveness category are positive responsiveness, resolvability, and perfect-tie responsiveness. These criteria are all concerned with ensuring that small changes that seem like they should break ties actually do so in expected ways.

The criteria in the stability category are monotonicity, homogeneity, combinativity, and independence of clones. These criteria are all concerned with

ensuring that the outcome of an RBVPR does not change as a result of a change in input that arguably should not change it.

The criteria in the qualitative attractiveness category are lucidity, ease of use, cost of computation, and resistance to strategy. These criteria can be possessed to a greater or lesser degree and are susceptible to argument about the extent to which vote processing rules possess them.

The following implications among properties have been noted:

mutual majority consistency ⇒ majority consistency
Schwartz consistency ⇒ Smith consistency ⇒ Condorcet consistency ⇒ majority consistency
positive responsiveness ⇒ monotonicity
combinativity ⇒ homogeneity
two out of three of Condorcet consistency, perfect-tie responsiveness, and combinativity ⇒ the absence of the third

The implications listed above will be used in the next chapter, where the criteria developed in this chapter are used to evaluate vote-processing rules.

Notes

1 See Chapter 10 and Arrow, *Social Choice and Individual Values* (New Haven, 1963), pp. 2, 6–7, 15.
2 See Richard Zeckhauser, "Voting Systems, Honest Preferences and Pareto Optimality," *American Political Science Review* 67 (1973): 934–46; Bruce Ackerman, *Social Justice in the Liberal State* (New Haven, 1980), pp. 285–9.
3 Condorcet, *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix* (Paris, 1785).
4 John H. Smith, "Aggregation of preferences with variable electorate," *Econometrica* 41/6 (1973): 1027-41, p. 1038.
5 Thomas Schwartz has called this criterion GETCHA, for GEneralized Top CHoice Assumption, *The Logic of Collective Choice* (New York, 1986), p. 141.
6 Ibid., p. 142.
7 Invariant loss consistency is sometimes called Condorcet losing consistency.
8 H. P. Young, "Extending Condorcet's Rule," *Journal of Economic Theory* 16 (1977): 335–53.
9 I have previously given the name "additivity" to the property of combinativity. Smith ("Aggregation of preferences with variable electorate," p. 1029) called it separability. Others have called it consistency.
10 At the web site http://condorcet.org/emr/defn.shtml, Blake Cretney defines "bury" as "Insincerely ranking an alternative lower in the hope of defeating it." For Cretney, the opposite of "bury" is "compromise."

This page intentionally left blank

Chapter 13

Vote-Processing Rules for Selecting One Option from Many When Votes Have Predetermined Weights: Alternatives to Plurality

Introduction

This chapter provides evaluations of a variety of vote-processing rules by the criteria that were developed in Chapter 12 and displayed in Table 12.2. The rules are considered in roughly historical order, except that rules that were not designed as ranking-based rules are considered first. Many of the rules are too complex to be described in this introductory section, so each will be described as it arises. The chapter closes with a table that displays the properties of all the rules that are considered and a discussion of the factors that could reasonably lead to a decision that one rule or another is best for a particular situation.

Plurality

The most common rule for electing one option from many is the plurality rule: Every voter names one nominated option as the one for whom he or she votes, and the option with the largest number of votes is the winner.

Plurality is an extension of majority rule to elections for more than two options in the simplest and most natural way. Under majority rule for elections between pairs of options, the option with the great*er* number of votes is the winner, and under plurality for multiple options, the option with the great*est* number of votes is the winner. Our long history of using plurality gives it an authority that would be hard for any other rule to match.

People often refer to the plurality rule as "majority rule," but this designation is inappropriate. "Majority" means "more than half," and when there are more than two options, the option with the greatest number of votes may have less than half.

Compared to other RBVPRs, the customary plurality rule uses rankings with limited domains. In effect, voters are required to report all options but one as tied for second place in their rankings. However, it is possible to give plurality a universal domain by having voters report their full rankings of options and using

only the information about which option each voter ranks first to count the votes. These two versions of plurality are considered in the two subsections that follow.

Plurality as a Function of Rankings with Limited Domains

When plurality is interpreted as an RBVPR with a domain limited to rankings in which all options but one are tied for second place, this limited domain is its principal limitation. It is unclear whether plurality as a function of rankings with limited domains should be said to possess mutual majority consistency, because when all options but one are tied for second place, it is not possible for two or more options to be ranked ahead of all others by any voters. Apart from universal domain and mutual majority consistency, plurality as a function of rankings with limited domains possesses all of the non-qualitative properties that are used as criteria in this chapter: majority consistency, Condorcet consistency, Smith consistency, Schwartz consistency, invariant loss consistency, positive responsiveness, resolvability, perfect-tie responsiveness, monotonicity, homogeneity, combinativity, and independence of clones.

Plurality as a function of rankings with limited domains possesses Schwartz consistency (and therefore Smith consistency, Condorcet consistency, and majority consistency) because the plurality winners are the only elements of the GOCHA set. It possesses invariant loss consistency because an option that loses to all others in paired comparisons has a lower vote total than any other option. It possesses positive responsiveness, and therefore monotonicity, because a rise in the position of a tied winner from tied for second place to first place in one voter's ranking will resolve the tie in favor of the option that rises. It possesses resolvability because, if there is a tie between option x and one or more other options, then adding another vote that ranks x first will cause x to win. For a similar reason, it possesses perfect-tie responsiveness. Plurality as a function of rankings with limited domains possesses combinativity because any inequalities in the plurality scores of options that are repeated in two batches of rankings will be preserved when the two batches are combined. Since combinativity implies homogeneity, it possesses homogeneity as well. When there are more than two options, a set of options can be clones under plurality as a function of rankings with limited domains only if none of them receive any votes. If there are three or more options and option x receives a vote, then no option y can be a clone of x because a voter who votes for x is thereby ranking y as tied with all options other than x and y, which is inconsistent with y being a clone of x. Under plurality as a function of rankings with limited domains, all options that receive no votes are clones, and with respect to such options plurality is independent of clones.

With respect to measures of qualitative attractiveness, plurality as a function of rankings with limited domains is as lucid, as easy to use, and as simple to count as a voting rule could be. Therefore I give plurality as a function of rankings with limited domains scores of 10 for lucidity, ease of use, and cost of computation.

Susceptibility to strategy is a significant weakness of plurality. It is not susceptible to a burying strategy because no distinctions are made among options that a voter does not rank first. The indirect hoisting strategy has no applicability to a rule such as plurality that records each option's performance as a single score. The strategy to which plurality is susceptible is the direct hoisting strategy.

The option for which a voter votes under plurality is not referred to as the voter's first choice because it is widely accepted that in elections by plurality, voters often do not cast their votes for their first choices. Voters who believe that their first choices are very unlikely to come close to winning may understandably decide to vote not for their favorites among all nominated options, but rather for their favorites among the options that seem to have reasonable prospects.

For each of the 87 elections discussed in chapter 9, Figure 13.1 shows, for the strategy requiring the fewest participants (or strategants, as I shall call them henceforward), a scatter plot of the percentage of voters who must be strategants for the strategy to succeed, plotted against the percentage of the electorate whose interests would be advanced by the strategy. For example, if there were an election with 200 voters, with 80 votes for the first-place finisher, 71 votes for the second-place finisher, and 49 for other options, a voter who favored the second-place finisher would need to find 10 voters, or 5 per cent of the electorate, to switch their votes from other options to the second-place finisher to cause the second-place finisher to win. If 20 of the 49 voters who favored other options reported rankings that showed the second-place finisher ahead of the first-place finisher, then the percentage of the voters who were available for the strategy would be 10 per cent.

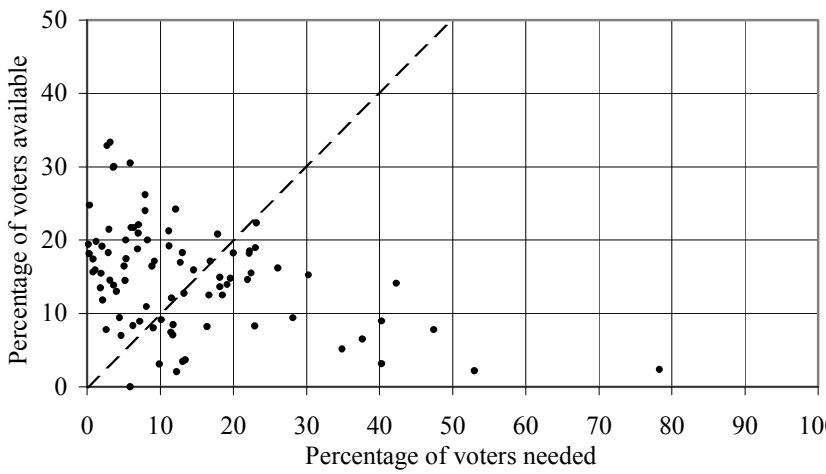


Figure 13.1 Scatter plot of the percentages of voters needed and available for minimal strategies under plurality, for 87 elections

For the 37 elections represented by dots below the 45-degree line in Figure 13.1, the percentage of voters who are needed as strategants is greater than the percentage available, and the election is not susceptible to strategy. For the 50 elections represented by dots above the 45-degree line, a strategy is feasible, but it requires more strategants when the dot is farther to the right. Furthermore, if the dot is not so far to the right, but it is close to the 45-degree line, then the strategants are harder to find. A measure of the difficulty of putting together a successful strategy that combines the percentage of voters who must participate and the percentage of the voters whose interests would be served by the strategy is the expected value of the percentage of the voters who would need to be contacted before one would encounter the needed participants, assuming no knowledge of which voters would be more likely to find the strategy attractive. This measure of the difficulty of forming a strategic coalition can be computed as

$$D = \sum_{T=0}^{U} \left[\frac{T+N}{V} \frac{\dfrac{A!}{(A-N)!}\dfrac{U!}{(U-T)!}}{\dfrac{V!}{(V-N-T)!}} \frac{(T-N-1)!}{T!(N-1)!} \right], \tag{13.1}$$

where T is the number of voters who turn the organizer down when contacted because their interests are not served by the strategy, U is the total number of voters who are unavailable because their interests are not served by the strategy, N is the number of strategants needed, A is the number of voters whose rankings of the options make them available as strategants, and V is the total number of voters. In this expression, the first fraction is a percentage of the voters that one might need to contact, the second fraction is the probability of encountering, in one particular order, the combination of voters who are and are not benefited by the strategy that the first fraction specifies, and the third fraction is the number of different orders in which that particular combination of persons who are and are not benefited might be contacted.

The difficulty of forming a strategic coalition, D, is used to compute R, which measures resistance to strategy on a scale of 0 to 10. R is defined as the average over elections of $10D$ for elections are susceptible to strategy and 10 for elections that are not. For the plurality rule, R (computed from the 87 elections in the sample) is 6.3, and this is the score that I give plurality for resistance to strategy.

One consequence of the susceptibility of plurality to the direct hoisting strategy is that voters tend to organize themselves into just two parties when the plurality rule is used. Each party decides which one option to offer to the electorate, so that similar options will not "split the vote," resulting in the election of the other party's option. This organization of the electorate into two parties has advantages and disadvantages. One advantage is that the winning option is likely to have the authority that comes from receiving a majority of the votes. Another advantage is

that a voter who wants to spend minimal time learning about options can rely on parties to give him just two options about which to learn. The corresponding disadvantage is that the electorate as a whole will not have the opportunity to choose among as many options as would be available under other rules.

The tendency of plurality to foster direct hoisting and just two parties also means that if there is a third perspective that would be more attractive to voters than those offered by the two principal parties, it may languish indefinitely as a consequence of voters' strategic calculations. Voters may decide that since so few other voters will be voting for the third alternative, there is no point in giving their votes to it. This seems to have been the fate of the Liberal Party in many constituencies in England.

In sum, plurality as a function of rankings with limited domains is simple, natural and familiar. But its susceptibility to vote splitting fosters strategic voting that makes it sometimes questionable whether the winning option can be described as the one that ought to have won.

Plurality as a Function of a Universal Domain

One can question whether plurality should be described as a function of rankings with limited domains. It would be possible to have voters report their full rankings of the options and then select as the winning option the one that was ranked first by the most voters. Under this interpretation, plurality possesses universal domain, but it lacks mutual majority consistency, Condorcet consistency, Smith consistency, Schwartz consistency, invariant loss consistency, positive responsiveness, and independence of clones. The proofs that it possesses the other properties are essentially the same as the proofs that plurality as a function of rankings with limited domains possesses those properties.

Example 13.1 shows that plurality as a function of a universal domain lacks mutual majority consistency.

Example 13.1

8	7	6
x	*y*	*z*
z	*z*	*y*
y	*x*	*x*

(When, as in this case, an example is more complex than is necessary to demonstrate the point for which it is introduced, it is either because of the historical significance of the example or because the same example will be used later to demonstrate other points.) Options *y* and *z* have a mutual majority in Example 13.1, but *x* is the plurality winner.

Example 13.1 also shows that plurality as a function of a universal domain is not Condorcet consistent (and is therefore neither Smith consistent nor Schwartz consistent). In paired comparisons from Example 13.1, *z* beats *x* by 13 votes to 8 and beats *y* by 14 votes to 7, and is therefore dominant, but *x* is the plurality winner. Example 13.1 further shows that plurality as a function of a universal domain lacks invariant loss consistency, since *x* is the plurality winner and loses to both *y* and *z* in paired comparisons,

Plurality as a function of a universal domain lacks positive responsiveness because an incremental upward movement to any position other than the top of a ranking will have no effect on the result.

Example 13.1 shows as well that plurality as a function of a universal domain is not independent of clones. Options *y* and *z* are clones in Example 13.1, and the deletion of option *y* changes the winner from *x* to *z*.

I give plurality as a function of a universal domain a score of 6 for lucidity, because people might be confused by the request for a full ranking, when only the information about which option was ranked first would be used. I give it a score of 5 for ease of use, to reflect the cost of providing a full ranking. The score for cost of computation is reduced from 10 to 9, to reflect the cost of having more information than is necessary. The score for resistance to strategy remains 6.3, since both versions of plurality offer the same strategic opportunities.

Approval Voting

Approval voting is a variation on plurality in which each voter is permitted to give one vote to as many of the options as he or she chooses. Several persons invented it independently in the mid 1970s.[1] Approval voting may be considered either as a function of rankings with limited domains or as a function of the combination of rankings with unlimited domains and cut-off points.

Approval Voting as a Function of Rankings with Limited Domains

As a function of rankings with limited domains, approval voting is an RBVPR that requires voters to divide the options into a set that is tied for first place (those to which they give a vote) and a set that is tied for last place (those to which they do not give a vote).

It cannot be expected that under approval voting voters will invariably prefer all options to which they give a vote to all those to which they do not give a vote. For example, if there is an election by approval voting in which two of the options are Republicans and two are Democrats, and voters expect that most other voters will vote either for both Republicans or for both Democrats, while being unsure whether there are more Republicans or Democrats among the electorate, then, whether they are Republicans or Democrats, they will have an incentive to vote for

their preferred Republican and their preferred Democrat, to maximize the probability of affecting the outcome in a direction that they prefer. While the set of options receiving a voter's votes is thus not uniformly preferred to those not receiving his votes, rational voting will nevertheless require a voter under approval voting to give a vote to the option he or she favors most and not to give a vote to the option he or she finds least attractive.

As a function of rankings with limited domains, approval voting possesses all of the non-qualitative properties other than universal domain that are tabulated in this chapter. It is Schwartz consistent because cycles are not possible, so that the approval winners are the sole elements of GOCHA set. It is therefore also Smith consistent, Condorcet consistent, and majority consistent. It is mutual majority consistent because if there is a set of options that is ranked ahead all others by a majority of voters, then those that are ranked ahead receive an approval vote from a majority while the others do not, so that none of the others could possibly be an approval winner. Approval voting as a function of rankings with limited domains possesses invariant loss consistency because an option that loses to even one other option in a paired comparison cannot be the approval winner.

Approval voting as a function of rankings with limited domains is positively responsive, and therefore monotonic, because a tied winner that rises in one voter's ranking thereby receives an additional vote, resolving the tie in favor of that option. It is resolvable because the addition of a vote with a tied winner ranked solely first resolves the tie in favor of that option. For a similar reason, it is perfect-tie responsive. It possesses combinativity, and therefore homogeneity, because the vote totals from combining two batches of rankings will be the sum of the vote totals of the separate batches. It possesses independence of clones because the removal of a clone has no effect on the vote totals of the remaining options, and therefore no effect on their winning status.

Approval voting is very easy to understand: Vote for as many options as you like. I give it a score of 9 for lucidity, just below plurality. Deciding which options to give a vote to is easier than deciding on a ranking of the options but noticeably harder than deciding which one option to vote for, so I give approval voting as a function of rankings with limited domains a score of 7 for ease of use. Determining the winner is nearly as easy as with plurality, so I give it a score of 9 for cost of computation.

The fact that approval voting requires voters to divide options into just two categories is the central limitation of approval voting. It is virtually impossible to do this without thinking strategically. It cannot be expected that voters will simply ask themselves whether they "approve" of options. Suppose a voter was offered the chance to give an approval vote to any of ten options and gave a vote to four. It cannot be expected that if just the four options to which the voter gave a vote were offered, the voter would have given a vote to all of them, nor that if just the six that were not given a vote were available, the voter would have voted for none of them. Thus receiving a vote from a voter under approval voting should not be

interpreted as meaning that the voter "approves" of the option. Receiving a vote means only that when given the chance to divide the options into a group that receives a vote and a group that does not, the voter chooses to put the option in the group that receives a vote. Since a voter cannot be expected to vote simply by asking, "Is this an option of which I approve?" it must be expected that voters under approval voting would be asking themselves, "Is this an option to which I wish to give a vote, considering the likely consequences of doing so?" It is unfortunate to put voters in this position, as it means they must confront such questions as, "Shall I give a vote to all of the options that I believe are adequate, thus helping to insure that the winner will be adequate, or shall I trust other voters to do that and make a distinction among the adequate options?"

The general strategy that works under approval voting, with reasonable foresight, is to identify the two options that are most likely to win and give only one of them a vote. This involves "burying" the less attractive option if one had been inclined to vote for it and "directly hoisting" the more attractive option if one had been inclined not to vote for it.

To estimate the susceptibility of approval voting to strategy from the available data, I assume that a voter would give an approval vote to every option mentioned on his or her ballot, unless all options are mentioned, in which case the last is dropped. The number of votes needed for the minimal strategy to succeed is one more than the difference between the top two approval scores. The number of voters who are available to implement the strategy is the number of voters who gave votes to both options and ranked the option with fewer approval votes ahead of the option with more approval votes. Thus the strategy that I measure is a burying strategy. Some voters would presumably be interested in a direct hoisting strategy, under which they gave a vote to an option of which they did not approve, because it was better than the one that was perceived as likely to win. But I have no way of knowing the ranking of options that are not mentioned on ballots, so I must leave this possibility unexplored.

Figure 13.2 shows a scatter plot, for the 87 elections in the sample, of the percentage of voters that must be strategants for the minimal strategy to succeed under approval voting, plotted against the percentage of the electorate whose interests would be advanced by the strategy, given the rankings that were reported in that election. Where there were 37 elections that were not susceptible to strategy under plurality, there are just 15 under approval voting. The resistance to strategy falls from 6.3 for plurality to 3.9 for approval voting. (To limit the size of the diagram, two instances where the percentages of voters who benefit exceed 50 per cent are shown as 50 per cent.)

Approval Voting as a Function of Rankings Plus Cut-off Points

If a voter submits a ranking in which one or more options are ranked above at least one option and below at least one other, then it is not possible to turn that ranking

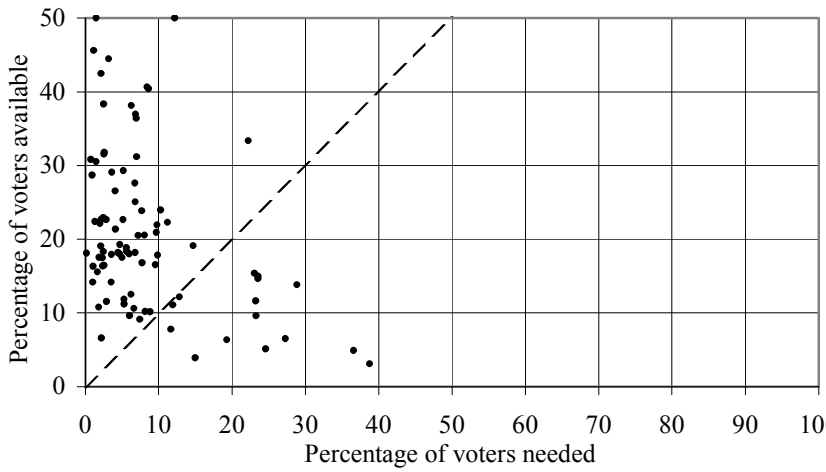


Figure 13.2 Scatter plot of the percentages of voters needed and available for minimal strategies under approval voting, for 87 elections

into an approval vote without additional information. One needs a cut-off point to separate the options that receive a vote from those that do not. Thus if one wishes to express approval voting as a function of rankings with a universal domain, one must express it as a function of the combination of a ranking and a place in the ranking that separates the options that receive a vote from those that do not. If approval voting is so construed, it lacks majority consistency and also all of the properties that plurality as a function of rankings lacks: mutual majority consistency, Condorcet consistency, Smith consistency, Schwartz consistency, invariant loss consistency, positive responsiveness, and independence of clones. The proofs that it possesses the other properties are essentially the same as the proofs that approval as a function of rankings with limited domains possesses those properties. Example 13.2 shows that approval as a function of rankings plus cut-off points lacks majority consistency. The horizontal lines separate the options that receive approval votes from those who do not. Option *x* has a majority of the first-place votes, but option *y* is the approval-voting winner.

Example 13.2

2	1
x	*y*
y	*x*
z	*z*

Approval voting as a function of rankings plus cut-off points lacks positive responsiveness because an upward shift in the position of a tied winner in one

voter's ranking, among those receiving a vote or among those not receiving a vote, will not affect any option's vote total. When Example 13.1 is interpreted as an example of approval voting in which each voter chooses to vote for just one option, it shows that approval voting lacks all of the other properties that plurality as a function of rankings lacks.

While approval voting formally lacks independence of clones, there is a sense in which it resolves the problem of clones. The concept of clones was developed to deal with situations in which options are virtually identical. And if some options are virtually identical, then each voter under approval voting is likely to vote for either all of them or none of them. In this case, the elimination of a clone will have no effect on the winning status of any other option. Therefore, for approval voting, it is sensible to add to the definition of clones the requirement that every voter either votes for all options in the set or votes for no option in the set. Under this definition, approval voting is independent of clones.

With respect to qualitative attractiveness, I give approval voting as a function of rankings plus cut-off points a score of 6 for lucidity, because the unused ranking information can be confusing. I give it a score of 4 for ease of use because of the cost of determining a ranking plus a cut-off point. I lower the cost of computation from 9 to 8, to reflect the burden of unnecessary information. The strategic opportunities are the same under either version of approval voting, so the score for resistance to strategy remains 3.9.

Range Voting

Range voting is the practice of having each voter assign each option a score in a specified range and declaring the winner to be the option with the highest average score. Like plurality and approval voting, range voting asks voters to provide information other than rankings of the options. Unlike plurality and approval voting, range voting asks voters for information from which rankings can be deduced. This makes it easier to specify whether range voting possesses the properties by which ranking-based vote processing rules are evaluated, although redefinitions of concepts are needed in some cases.

Range voting possesses none of the properties in the consistency category. When concepts are redefined to reflect the domain of the rule, it possesses all other properties.

For range voting, an incremental upward movement for an option is an increase in the score assigned to the option by one voter. With this definition, range voting possesses positive responsiveness (and therefore monotonicity) because the incremental upward movement breaks a tie and produces a win for the option that moved upward. Range voting possesses resolvability and perfect tie responsiveness because the addition of a vote that ranks an option first always breaks a tie and produces a win for that option. It possesses combinativity (and

therefore homogeneity) because the scores of options for any combination of ballots will be weighted averages of their scores of the separate parts, and taking a weighted average preserves inequalities. For range voting, a set of options should be considered clones if for each voter there is a score that that voter assigns to all options in the set. For this definition, range voting is independent of clones because the removal of an option has no effect on the score of any other option.

Range voting and all other rules that are discussed later in this chapter possess universal domain by the nature of the input that they require from voters. Since this reason is the same for all of these rules, it will not be mentioned in subsequent discussions.

Example 13.3 shows that range voting lacks majority consistency and invariant loss consistency. In this example, three voters rank three options on a ten-point scale. Option *x* has a majority of the first-place votes, but option *z*, which loses to both of the other options in paired comparisons by majority rule, has the highest average score. Since range voting lacks majority consistency, it also lacks mutual majority consistency, Condorcet consistency, Smith consistency, and Schwartz consistency.

Example 13.3

2	1
x: 5	*z*: 9
y: 4	*x*: 2
z: 3	*y*: 1

Range voting is extremely easy to understand. I give it a score of 10 for lucidity. Its ease of use is comparable to that of rules that require a ranking, so I give it a score of 5 in that dimension. The cost of computation is only slightly greater than that of approval voting, so I give it a score of 8 in that dimension.

Range voting is highly susceptible to strategy, both in the form of burying a disfavored contending option by giving it a minimal score and in the form of hoisting a favored contending option to the top of the range. When these strategies are applied to their limits, all voters give all options scores that are either at the top or the bottom of the permitted range, and range voting becomes equivalent to approval voting.

To measure the susceptibility of range voting to strategy, rankings must be turned into range votes. I assumed a range of 0 to 100, with each voter's range votes equally spaced in the range. Thus if a voter ranked just two options, they were assumed to have been given range votes of 75 and 25, and all other options were assumed to have been given range votes of 0. The strategy that was then measured was for voters who ranked the option with the second highest average score ahead of the option with the highest average score to give the former a score of 100 and the latter a score of 0. This strategy can be characterized as a

combination of a direct hoisting strategy and a burying strategy. The number of voters available to implement the strategy is the number of voters who ranked the option with the second highest average score ahead of the option with the highest average score. The number of votes need for the strategy to succeed is the number of votes that are available multiplied by the ratio of the points needed to the points available.

Figure 13.3 shows a scatter plot, for the 87 elections in the sample, of the percentages of voters available and needed for the strategy examined to succeed. There are just 12 elections that are not susceptible to strategy. The resistance to strategy of range voting is 4.0.

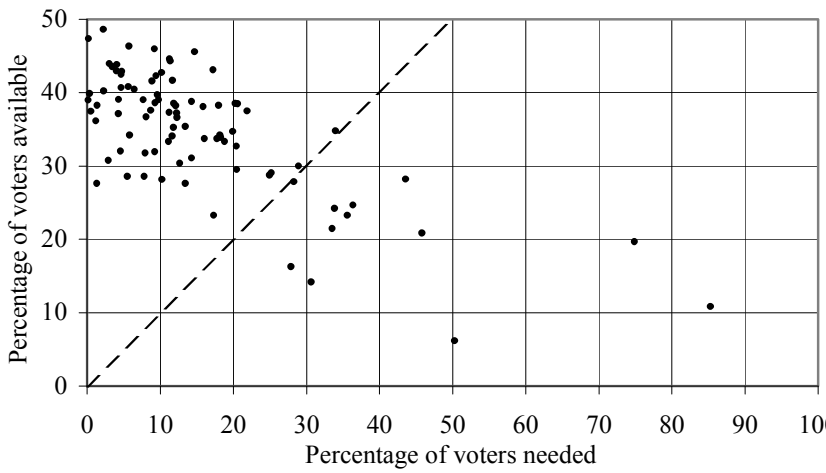


Figure 13.3 Scatter plot of the percentages of voters needed and available for minimal strategies under range voting, for 87 elections

The Borda Rule

Systematic investigation of alternatives to plurality appears to have begun with the work of Jean-Charles de Borda, a French military officer and scientist of the eighteenth century.[2] Borda read a paper, "Sur la Forme des Élections," to the French Academy of Sciences on June 16, 1770. There is no surviving record of that paper. (Four members were charged with reporting on it, but they apparently did not do so.[3]) Fourteen years later, four days after the Academy heard a report on Condorcet's ideas (discussed in the next section), which were then in manuscript, Borda read another paper to the Academy, "Mémoire sur les Élections au Scrutin." In a footnote to the published version of this paper, he says that the ideas in this paper had already been presented to the academy fourteen years

earlier. Therefore it is possible to be reasonably confident that the ideas in the surviving paper are close to those of the original paper.[4] Borda considers (in literary rather than tabular form) the example given earlier as Example 13.1, but with the options ranked second and third by the voters in the first column not specified.[5] He mentions that on a single vote x will win, even though the electorate would prefer either y or z to x by a margin of 13 to 8. This suggests that Borda was concerned with the consequences of vote splitting under the plurality rule and had not formulated the criterion of Condorcet consistency.

Borda goes on to suggest that the estimate of the merit that any voter attaches to any given option should increase by the same amount for each alternative option above which the given option is ranked.[6] Thus options should be scored by a system in which an option receives 1 point for being in second to last place, 2 points for being in third to last place, and so on up to $r - 1$ points for first place, where r is the number of options that are ranked. He mentions that the result is the same as that which occurs if all paired comparisons of options are made and an option receives one point for each vote it receives in each paired comparison.

This alternative specification of Borda's rule suggests the idea of a *matrix of paired comparisons*. In such a matrix, the entry in row i, column j is the number of voters who rank option i ahead of option j. Suppose that the votes are those given by Example 13.4.

Example 13.4

9	8	15	16
v	*w*	*x*	*y*
w	*v*	*z*	*z*
x	*x*	*y*	*v*
y	*z*	*w*	*w*
z	*y*	*v*	*x*

The matrix of paired comparisons for Example 13.4 is

	v	*w*	*x*	*y*	*z*
v	0	25	33	17	17
w	23	0	33	17	17
x	15	15	0	32	32
y	31	31	16	0	25
z	31	31	16	23	0.

As an example of how the numbers are calculated, the 25 in the second column of the first row derives from the 9 voters in the first column of the example plus the 16 voters in the last column. These are all of the voters who rank v ahead of w. If the matrix of paired comparisons has been calculated, the Borda score for any

option can be computed as the sum of the numbers in the option's row of this matrix. Thus the options' Borda scores for Example 13.4 are *v*: 92, *w*: 90, *x*: 94, *y*: 103, *z*: 101, and *y* is the winner by the Borda rule. As a check on one's calculations, it is worth noting that if there are N voters and r options, and no ties in any ranking, then the sum of the Borda scores is $Nr(r-1)/2$. Thus with 48 voters and 5 options in Example 13.4, the sum of the Borda scores is 480.

Duncan Black pointed out that the Borda rule makes no provision for the possibility that a voter will want to rank two or more options as tied. He suggested that the natural way to allow for ties in rankings is to take the usual Borda score and then subtract one point from each option's score for each alternate option that is ranked above it by any voter.[7] A score calculated in this way will be called an adjusted Borda score.

One way of computing adjusted Borda scores is by the use of a *matrix of majorities*. In this matrix, the entry in row i, column j is the number of voters who rank option i ahead of option j, minus the number who rank option j ahead of option i. That is, the matrix of majorities is the matrix of paired comparisons minus its transpose. For Example 13.4, the matrix of majorities is

	v	*w*	*x*	*y*	*z*
v	0	2	18	–14	–14
w	–2	0	18	–14	–14
x	–18	–18	0	16	16
y	14	14	–16	0	2
z	14	14	–16	–2	0.

The notation $M(i, j)$ will be used to denote the majority of i over j, that is, component ij of the matrix of majorities.

The adjusted Borda score for an option can be calculated either as the sum of the option's row of the matrix of majorities, or as the difference between its row and column sums in the matrix of paired comparisons. Thus for Example 13.4, the adjusted Borda scores of the options are *v*: – 8, *w*: – 12, *x*: – 4, *y*: 14, *z*: 10, and *y* is again the winner.

The mathematical relationship between an option's Borda score and its adjusted Borda score, when no voters have ranked any options as tied, can be derived from the fact that the total number of options below which an option could have been ranked is $N(r-1)$, where N and r are again the numbers of voters and options respectively. An option with a Borda score of B is ranked below $N(r-1)-B$ options and therefore has an adjusted Borda score of $A = 2B - N(r-1)$. Since the adjusted Borda score is a linear function of the Borda score, both scores yield the same order of the options when no voter has ranked any options as tied. In that case, solving the expression above for B, the Borda scores can be computed from the adjusted Borda scores by $B = [A + N(r-1)]/2$. Because of the anti-symmetry of the matrix of majorities, the sum of the adjusted Borda scores is always 0.

While the Borda rule deals reasonably well with the problem of vote splitting that arises under plurality, it still generates strong incentives for strategic voting. If there are many options but just two that seem reasonably likely to win, then a voter will be inclined to bury his less favored of the two at the bottom of his ranking and hoist the more favored to the top. As compared to placing the options in adjacent positions in one's ranking, placing them at the bottom and top increases the difference between their scores by $r - 2$ points. If a large number of voters behave in this way, it would be possible for a third option that was regarded as inferior to both leading contenders by most of the voters to obtain the highest Borda score. Thus the Borda rule is much less attractive as a rule for voters who are self-interested advocates than for voters who are presumed to be disinterested judges.

A generalization of the Borda rule that is often used is to assign some number of points to a first-place vote, a smaller number to a second-place vote, and so on. Such systems are sometimes used to rank U.S. college sports teams.

The Borda rule possesses universal domain, invariant loss consistency, positive responsiveness, resolvability, perfect-tie responsiveness, monotonicity, homogeneity, and combinativity, but not majority consistency, mutual majority consistency, Condorcet consistency, Smith consistency, Schwartz consistency, or independence of clones. The adjusted Borda rule possesses the same properties, differing only in its treatment of rankings containing ties.

These rules possess invariant loss consistency because an option that loses to all others will have a row of the matrix of majorities in which all non-diagonal components are negative, and therefore the sum is negative. Since the sum of all rows of the matrix of majorities is zero, there must be at least one row with a greater sum. They possess positive responsiveness (and therefore monotonicity) because any incremental upward movement of an option will increase that option's Borda score and no one else's, thus breaking any tie in which the option is involved, or maintaining its winning status if it already wins. They possess resolvability and perfect-tie responsiveness because any tie can be broken in favor of any tied option by adding a ranking that places that option first. They possess homogeneity and combinativity because the Borda scores from combining two or more batches of rankings are the sums of the Borda scores from the separate batches.

Eliminating the second column in Example 13.1 provides an example that shows that the Borda rule is not majority consistent. Then x has a majority of the votes, but the Borda scores of options x, y, and z are 16, 6, and 20 respectively, so that z is the Borda winner. Since the Borda rule is not majority consistent, it is not mutual majority consistent, Condorcet consistent, Smith consistent or Schwartz consistent.

That the Borda rule is not independent of clones can be seen by eliminating option v, a clone of w, from Example 13.4, resulting in adjusted Borda scores of –10, 14, 0 and –4 for options w, x, y and z respectively. Since the elimination of a clone changes the winner from y to x, the Borda rule is not independent of clones.

Turning to qualitative attractiveness, I give the Borda rule a score of 9 for lucidity, because this rule is the way of aggregating opinions about more than two options that comes most readily to people's minds when they would like something more sophisticated than plurality. It is given a score of 5 for ease of use, along with all rules based on rankings. It is given a score of 7 for cost of calculation because it is distinctly more difficult to calculate than plurality, but not nearly as difficult as some other methods. A number of other rules share with the Borda rule the quality that the result can be calculated quickly from the matrix of majorities. All these rules are given scores of 7 for cost of calculation.

To measure the susceptibility of the Borda rule to strategy, I calculated, for the 87 elections in the sample, the frequency with which the option that was ranked second by the Borda rule could be brought up to or ahead of the option ranked first, by having the voters who ranked the former ahead of the latter hoist the former to the tops of their rankings and bury the latter at the bottoms of their rankings. The number of voters available to implement the strategy is the number of voters who ranked the option with the second highest Borda score ahead of the option with the highest Borda score and did not vote in the strategically optimal way. The number of votes need for the strategy to succeed is the number of votes available multiplied by the ratio of the points needed to the points available.

Figure 13.4 shows a scatter plot, for the 87 elections in the sample, of the percentages of voters available and needed for the strategy examined to succeed. There are 20 elections that are not susceptible to strategy. The resistance to strategy of the Borda rule is 4.6.

Shortly after hearing Borda's 1784 paper, the French Academy of Science adopted his rule for deciding who would be elected to the Academy. They kept the

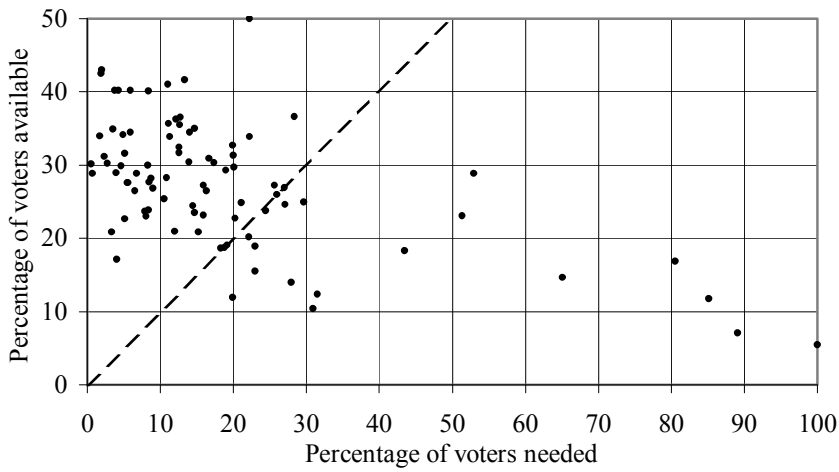


Figure 13.4 Scatter plot of the percentages of voters needed and available for minimal strategies under the Borda rule, for 87 elections

rule until 1800, when a new member, Napoleon Bonaparte, attacked it and succeeded in having it modified.[8]

A Digression on Examples

Since numerous examples will be considered, it is useful to note that for any antisymmetric matrix of integers, there is an example for which that matrix is the matrix of majorities. To see this, consider the combination of any ranking of options by one voter and the ranking that is obtained by reversing the order of all options except for the first and second, *i* and *j* respectively, which are ranked as tied at the bottom of the second ranking. The matrix of majorities produced by these two rankings consists of 0's everywhere except for component *ij*, which is 1, and component *ji*, which is –1. Replicating such a pair *N* times yields a matrix with *N* for component *ij*, *–N* for component *ji*, and 0's everywhere else. To produce any matrix of majorities, make a list of the row–column ordered pairs with positive components, and for each such pair with positive component *N*, include *N* voters who rank the row option first, the column option second, and all other options in any order, along with *N* other voters who have the opposite ranking except that the row and column options are tied at the bottom. That sequence of rankings will produce the chosen matrix of majorities.

The set of matrices of majorities that can be produced by strict rankings (rankings that do not include ties) is all matrices whose components are all even or all odd. Mixtures of odd and even integers cannot be produced because every addition of a vote either adds 1 to or subtracts 1 from every component of the matrix of majorities. To see that all other matrices can be produced, consider first the matrix of majorities for an election consisting of two rankings that have the options in opposite orders, except for the adjacent pair of options, *i* and *j*, that are in the same order in both rankings. The matrix will have 2 for component *ij*, –2 for component *ji* and 0's everywhere else. Replicating this pair or their opposites can produce any even integer in row *i*, column *j*, and adding other pairs of rankings can generate any desired anti-symmetric matrix of even integers. For a matrix of majorities consisting entirely of odd integers, first produce the matrix of even integers that is obtained by subtracting 1 from every component in the upper triangle, and then add one ranking consisting of the options in the order in which they appear in the rows and columns. The matrix of majorities for this ranking consists of 1's in the upper triangle, which is exactly what must be added to achieve the desired matrix. This shows that it is possible to use matrices of majorities for examples and be confident that there are examples of sequences of rankings that yield those matrices of majorities.

The Condorcet Rule

Although no formal record of Borda's 1770 paper was made, his ideas stimulated discussion in the French Academy of Science, which brought the subject of voting to the attention of Borda's colleague, the mathematician and political economist Marie Jean Antoine Nicolas Caritat, Marquis de Condorcet.[9] Borda's ideas and Condorcet's share the characteristic of approaching voting as a problem of estimation. Chapter 8 mentioned Condorcet's justification of majority rule as a way of estimating which of two options was more likely to be better when voters were assumed to be equally likely to make such determinations accurately. Condorcet approached the problem of voting over more than two options in a similar way. However, Condorcet's analysis is hard to understand. His work did not immediately initiate a tradition, and when scholars went back to his work generations later, they found some of what he said incomprehensible.[10] Condorcet was very clear that when there is a dominant option, then that option should win. But he was aware that sometimes, as in Example 13.4, there is no dominant option. (The lack of a dominant option can be discerned quickly from the matrix of majorities, by the fact that no row of this matrix consists entirely of positive numbers.) For 200 years no one seems to have understood what Condorcet was saying. But Peyton Young worked diligently through Condorcet's logic and identified a highly plausible interpretation of Condorcet's intention and the errors he made that led his readers astray.[11] Condorcet argued that the amount of evidence in favor of a particular *order of the options* could be measured by the total number of votes with respect to pairs that were consistent with that order, and that the option that ought to win is the one at the top of the order for which there is the greatest evidence. (To distinguish between a voter's ballot and a potential election result, the first will be called a ranking of the options and the second an order of the options.) Condorcet's idea was suggested independently by Kemeny in the twentieth century.[12] Thus the rule that I call the Condorcet rule is sometimes called the Kemeny rule or the Kemeny-Young rule.

To implement Condorcet's suggestion, one scores each order of the options according to the sum of the components of the upper triangle of the matrix of paired comparisons when the options are put in that order, and then declares as the winner the option at the top of the order with the highest score. The same result is also achieved by the simpler task of scoring each order of the options by the sum of the negative components of the upper triangle of the matrix of majorities that occurs when the options are placed in that order (less negative sums being better). This can be seen from the following argument.

Every pair of options contributes two numbers to all matrices of paired comparisons, one of which appears in the upper triangle. To every upper triangle of a matrix of majorities, the pair contributes a component with a magnitude equal to the difference between the two numbers that it contributes to the matrix of paired comparisons, but with a sign that varies among orders. If a pair contributes

equally to the upper triangles of the matrices of paired comparisons for two orders of the options, then it also contributes equally to the corresponding matrices of majorities. Therefore the exclusion of components of matrices of majorities that are positive for both orders has no effect on the relative scores of any pair of orders. Similarly, the inclusion of components that are negative for both orders has no effect on the relative scores of any pair of orders. If a pair contributes different numbers to the upper triangles of two matrices of paired comparisons, then the contribution of this pair to the difference in sums of the upper triangles of the matrices of paired comparisons for the two orders is the difference between these two numbers. But this is also what the pair contributes to the difference in the sums of the negative components of the upper triangles of the matrices of majorities. Thus the difference in scores of a pair of orders when they are scored by the sums of the upper triangles of the matrices of paired comparisons is the same as the differences in their scores when they are scored by the sums of the negative components of their matrices of majorities.

The ease of scoring orders by the sum of the negative components of the upper triangle of the matrix of majorities can be illustrated by considering Example 13.4. The score of the order *vwxyz* is the sum of the negative components of the upper triangle of the matrix as shown, namely –56.

The Condorcet rule possesses universal domain, majority consistency, mutual majority consistency, Condorcet consistency, Smith consistency, invariant loss consistency, resolvability, monotonicity, and homogeneity, but not Schwartz consistency, positive responsiveness, perfect-tie responsiveness, combinativity, or independence of clones.

The Condorcet rule possesses mutual majority consistency (and therefore majority consistency) by the following argument. Let S be a set of options that are all ranked ahead of all other options by a majority of the voters, and let r be an order of the options that places an option not is S first. Since the first component of r is not in S, somewhere in r there must be a pair of adjacent options, i and j, with i not in S and j in S. Compare r with the order r' obtained by interchanging the positions of i and j. The matrices of majorities for orders r and r' have all of the same components, except for the component comparing i and j, which is negative in r and positive in r'. Therefore the Condorcet score for r' must be greater than the Condorcet score for r. Therefore no order of the options that places an element not in S at the top can be the order with the greatest Condorcet score.

The Condorcet rule possesses Smith consistency and Condorcet consistency for a similar reason. If there is a set, S, of options that beat all other options in paired comparisons by majority rule, and r is a ranking that places an option not in S first, then there is an i that not in S that is immediately above a j that is in S, and interchanging i and j produces an order of the options with a greater Condorcet score, so that no order of the options that places an option not in S can have the greatest Condorcet score.

The Condorcet rule also possesses invariant loss consistency for a similar reason. If i is an option that loses to all other options by majority rule, then an order of the options that places i first has a lower Condorcet score than the order obtained by interchanging i and the option immediately below it.

The Condorcet rule possesses resolvability because, in the event of a tie between order r headed by option i and order s headed by option j, adding a vote corresponding to order r will increase every component of the upper triangle of the matrix of paired comparisons for order r, but only some of the components of the upper triangle of the matrix of paired comparisons for order s, thereby resolving the tie in favor of order r.

The Condorcet rule possesses monotonicity because moving any option i ahead of any other option j in one voter's ranking increases by 1 the Condorcet score of every order that places i ahead of j and lowers by 1 the Condorcet score of every order that places i behind j. Thus the Condorcet score of any order that places i first increases by as much as any order's score increases, and the change cannot cause i to lose when it would otherwise win.

The Condorcet rule possesses homogeneity because repeating a sequence of rankings k times multiplies all Condorcet scores by a factor of k.

Example 13.5 shows that the Condorcet rule lacks Schwartz consistency.[13]

Example 13.5

5	5	4	4	3	3	2	2	1	1
x	z	y	x	z	y	w	y	w	z
y	w	z	w	x	w	z	x	y	x
w	x	w	y	w	z	x	w	x	w
z	y	x	z	y	x	y	z	z	y

The matrix of paired comparisons for Example 13.5 is

	w	x	y	z
w	0	0	2	4
x	0	0	10	–6
y	–2	–10	0	8
z	–4	6	–8	0.

For this example, w and x are tied winners by the Condorcet rule, but the GOCHA set consists only of w.

Example 13.6 shows that the Condorcet rule lacks positive responsiveness.

Example 13.6

1	2	1
x	*y*	*z*
y	*x*	*x*
z	*z*	*y*

For this example there is a tie between orders *xyz* and *yxz*, which generate matrices of majorities of

	x	*y*	*z*
x	0	0	2
y	0	0	2
z	–2	–2	0

and

	y	*x*	*z*
y	0	0	2
x	0	0	2
z	–2	–2	0

respectively.

If the voter in the third column interchanges *x* and *z*, then the new matrices of majorities for these two orders are

	x	*y*	*z*
x	0	0	4
y	0	0	2
z	–4	–2	0

and

	y	*x*	*z*
y	0	0	2
x	0	0	4
z	–2	–4	0.

The election is still a tie between *x* and *z*. Since the incremental upward movement by *x* did not cause *x* to win when it was previously tied, the Condorcet rule lacks positive responsiveness.

Example 13.7 shows that the Condorcet rule lacks perfect-tie responsiveness.

Example 13.7

3	3	3	1
x	*y*	*z*	*x*
y	*z*	*x*	*z*
z	*x*	*y*	*y*

The voters in the first three columns Example 13.7 form a perfect three-way tie. Thus for the Condorcet rule to possess perfect-tie responsiveness, option *x* must win for this example. However, there is a tie for the order with the greatest Condorcet score between *xyz* and *zxy*, with matrices of majorities of

	x	*y*	*z*
x	0	4	–2
y	–4	0	2
z	2	–2	0

and

	z	*x*	*y*
z	0	2	–2
x	–2	0	4
y	2	–4	0

respectively. Thus the Condorcet rule lacks perfect-tie responsiveness.

Example 13.8 shows that the Condorcet rule lacks combinativity.

Example 13.8

Sequence **R**:			Sequence **R′**:	
7	5	4	5	4
x	*y*	*z*	*x*	*z*
y	*z*	*x*	*z*	*y*
z	*x*	*y*	*y*	*x*

The matrices of majorities for sequences **R**, **R′** and **R** + **R′** are:

Sequence **R**

	x	*y*	*z*
x	0	6	–2
y	–6	0	8
z	2	–8	0

Sequence **R′**

	x	*y*	*z*
x	0	1	1
y	–1	0	–9
z	–1	9	0

Sequence **R** + **R′**

	x	*y*	*z*
x	0	7	–1
y	–7	0	–1
z	1	1	0

Under the Condorcet rule, *x* wins for sequence **R** and for sequence **R′**, but *z* wins for sequence **R** + **R′**. Thus the Condorcet rule lacks combinativity. However, it is worth noting that if the Condorcet rule selects the same *order* of a given set of options for two sequences of rankings, then, because of the additivity of inequalities among sums of upper triangles of matrices of paired comparisons, the Condorcet rule will select the same order for the combination of the two sequences of rankings.

Example 13.4 shows that the Condorcet rule is not independent of clones. Recall that *y* is the winner for this example, and note that *v* and *w* are clones. When the clone *w* is eliminated, the order with the largest Condorcet score is *xyzv*, with a matrix of majorities of

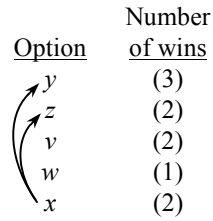
	x	*y*	*z*	*v*
x	0	16	16	–18
y	–16	0	2	14
z	–16	–2	0	14
v	18	–14	–14	0.

Since the elimination of *w*, a clone of *v*, changes the winner from *y* to *x*, the Condorcet rule is not independent of clones.

Turning to qualitative attractiveness, I give the Condorcet rule a score of 6 for lucidity, reflecting the complexity of the idea of ordering orders of the options. I give a score of 5 for ease of use, as with other rules requiring rankings.

The score of the Condorcet rule for cost of computation must reflect the difficulty of identifying the order of the options with the highest Condorcet score. While this can be quite difficult, there is a heuristic procedure for identifying the winner by the Condorcet rule that greatly simplifies the task. Order the options according to the number of other options they beat, and then move any option up if it is immediately below an option that it beats. Then use arrows to indicate instances in which an option beats another that is above it. Any ties can be indicated by dashed arrows. The results of applying this procedure to Example 13.4 are shown below.

Application of a heuristic procedure for the Condorcet rule to Example 13.4



If there are one or more places in such a list where one can draw a horizontal line without cutting any solid or dashed arrows (which does not occur in this case), then the winning option must lie above the highest such place, because the highest such line separates the Smith set from other options, and the Condorcet rule is Smith consistent. Furthermore, the winning option can be identified by finding the order of the Smith set with the best Condorcet score, ignoring all options outside the Smith set. This is because the winning order must place all of the Smith set ahead of all other options, and the overall order with the best score must begin with the Smith set in the order with its best score.

The second way in which the heuristic procedure is useful is in reducing the orders of the Smith set which must be considered. In the winning Condorcet order of the Smith set, every option after the first is ether below its predecessor in terms of this order, with no arrow joining them, or else it is above its predecessor and linked by an arrow. For the diagram shown, if an order starts at *w* or *x*, it must have *x* followed by the sequence *y*, *z*, *v*. There are two such orders. If an order starts at *v* it must follow *v* by either *w*, *x*, *y*, *z* or *x*, *y*, *z*, *w*. This adds two more orders. If an order starts at *z* it must proceed to *x*, stopping at any non-empty subset of *v* and *w* (three possibilities) after which it goes to *y* and whatever it

skipped on the first downward trip. An order starting at y can either put the options in the listed order (one possibility) or it can skip z and any non-empty subset of v and w (three possibilities) after which it goes to z and whatever it skipped on the first downward trip. Thus just 11 of the 120 possible orders of the options satisfy the requirement that every option beat the option immediately below it, and these 11 can be identified without excessive difficulty. Thus finding the Condorcet winner is substantially easier than it might at first seem.

For Example 13.4, the order with the best Condorcet score is the order that is shown in the illustration of the heuristic procedure, *yzvwx*, which yields the following matrix of majorities.

	y	z	v	w	x
y	0	2	14	14	–16
z	–2	0	14	14	–16
v	–14	–14	0	2	18
w	–14	–14	–2	0	18
x	16	16	–18	–18	0

The sum of the negative components of the upper triangle of this matrix is –32. Since this sum is greater (smaller in absolute value) than that for any other order of the options, y is the winner for Example 13.4 by Condorcet's rule. Although the heuristic procedure simplifies the identification of winners under the Condorcet rule, it is still quite tedious compared to other rules, and therefore the Condorcet rule is given a score of 3 for cost of computation. If matrices of majorities with highly involved cycles are encountered only rarely, this score may be unduly low.

The capacity of the Condorcet rule to resist strategy is strongly influenced by the fact that under the rule, an option that is dominant necessarily wins. No strategy of direct opposition can displace a dominant option, since for every challenger there is a majority that prefers the dominant option to the challenger. This implies that if there are no cycles, a successful strategy against the Condorcet rule must create a cycle. The rarity of cycles and the difficulty of creating them give the Condorcet rule a high resistance to strategy.

Suppose there is an election in which x is dominant, y is second, and z is third. The strategy against the Condorcet rule that sometimes works for such elections is for voters who prefer y to x to rank z ahead of x. If enough voters do this to cause z to beat x, and if the margin by which x beats y becomes the smallest margin, then the Condorcet rule will select y.

Suppose that honest voting yields the matrix

	x	y	z
x	0	a	b
y	$-a$	0	c
z	$-b$	$-c$	0,

with *a*, *b*, and *c* all positive. The strategy is to change *b* to a negative number that is greater in absolute value than *a*. Then number of vote changes needed to do this is $a + b + 1$, with a vote change counting for 2 if it comes from voter who had previously ranked *x* ahead of *z* and counting for 1 if comes from a voter who had previously ranked *x* and *y* as tied. But the strategy will not work at all if *a* is greater than *c*, because in that case the Condorcet rule selects the order *zxy* when the strategy is implemented, and the strategants do not agree that this is good.

For the 87 elections in the sample, it happened 51 times that *a* was greater than *c*, so that the strategy could not be implemented. Another 15 times there were not enough voters with preferences of the right sort to implement the strategy. Only 21 times was it possible to implement the strategy. Figure 13.5 shows a scatter plot, for the 87 elections in the sample, of the percentages of voters available and needed for the Condorcet strategy examined to succeed. For the elections where the strategy was not feasible, the percentage of voters needed is set arbitrarily at 100 per cent. The resistance to strategy of the Condorcet rule is 8.9.

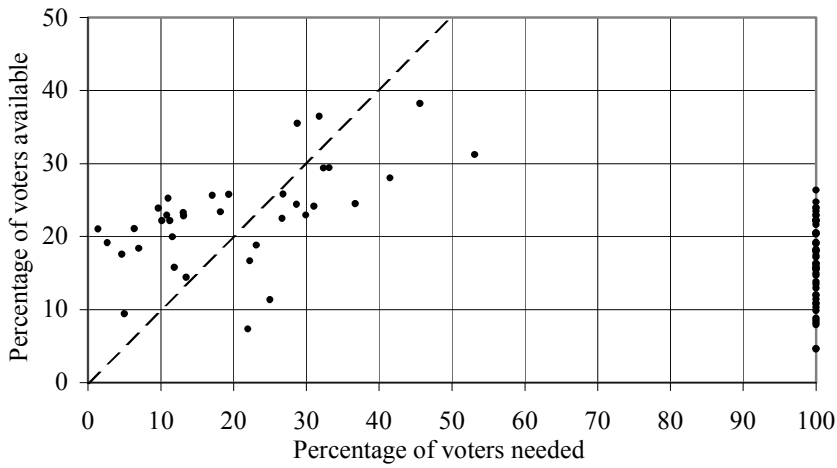


Figure 13.5 Scatter plot of the percentages of voters needed and available for minimal strategies under the Condorcet rule, for 87 elections

The fact that the Condorcet rule is not Schwartz consistent may seem to be a disadvantage for the rule. However, one can equally use the example that was used to show that the Condorcet rule is not Schwartz consistent (Example 13.5) to argue against Schwartz consistency. Is it utterly clear in this example that *w* is the only plausibly valid winner? If *w* and *x* had been the only options, then the result would clearly have been a tie. When *y* and *z* are added, it is clear that *w* should be ranked ahead of them, while *x* gets involved in a cycle with them. But if the cycle is broken with *x* ranked ahead of *y* and *z*, then it is not inconsistent to say that *x*

continues to be ranked as tied with *w*. Perhaps the criterion of Schwartz consistency is not compelling.

Two-Ballot Majority

Two-ballot majority (known also by other names such as majority run-off and two-round run-off) is a variation on plurality in which, if the option with the most votes does not have a majority of the votes, then a second round of voting is held between the top two options.[14] As with plurality, there is ambiguity about whether two-ballot majority is an RBVPR, since it does not require voters to rank the options, but the outcome can be determined from rankings.

Two-ballot majority, when interpreted as an RBVPR, possesses majority consistency, invariant loss consistency, and homogeneity, but not mutual majority consistency, Condorcet consistency (and therefore neither Smith consistency nor Schwartz consistency), monotonicity (nor therefore positive responsiveness), resolvability, perfect-tie responsiveness, combinativity, or independence of clones.

Two-ballot majority possesses majority consistency because an option that is ranked first by a majority is elected on the first ballot. It possesses invariant loss consistency because an option that loses to every other option in paired comparisons cannot possibly win in the second round. It possesses homogeneity because the vote totals from repeating a sequence of rankings *k* times will be *k* times the vote totals of the original sequence.

Example 13.9 shows that two-ballot majority lacks mutual majority consistency.

Example 13.9

4	4	4	5	6
x	*y*	*z*	*v*	*w*
y	*z*	*x*	*w*	*v*
z	*x*	*y*	*x*	*z*
w	*w*	*w*	*y*	*y*
v	*v*	*v*	*z*	*x*

The set $\{x, y, z\}$ has a mutual majority, but the two-ballot majority rule chooses *w*.

Example 13.1 shows that two-ballot majority lacks Condorcet consistency (and therefore Smith consistency and Schwartz consistency). Option *z* is dominant, but under two-ballot majority *z* is eliminated in the first round.

Example 13.10 shows that two-ballot majority lacks monotonicity.

Example 13.10

3	4	3	1
x	*y*	*z*	*z*
y	*z*	*x*	*y*
z	*x*	*y*	*x*

Counting this example by two-ballot majority rule, *x* is eliminated in the first round and *y* wins in the second. But if the voter in the last column were to move *y* ahead of *z*, then *x* and *z* would tie for fewest votes. If the tie were resolved in favor of *x*, then *x* would beat *y* in the second round of voting, so that *y*'s incremental upward movement costs it the election. Thus two-ballot majority lacks monotonicity, and therefore also lacks positive responsiveness.

Example 13.11 shows that two-ballot majority rule lacks resolvability.

Example 13.11

1	1	2	3
x	*x*	*y*	*z*
y	*z*	*x*	*x*
z	*y*	*z*	*y*

By two-ballot majority rule, in the first round there is a tie for second place between *x* and *y*. If *x* is eliminated, then *z* wins. If *y* is eliminated, then *x* wins. To guarantee that *z* will win, *x* must be definitely eliminated, by adding a vote that places *y* first. But if that is done, there will be a tie in the second round between *y* and *z*. So there is no way to add a vote that will break the tie in such a way as to guarantee that *z* will win. Thus two-ballot majority rule lacks resolvability.

Example 13.7 shows that two-ballot majority lacks perfect-tie responsiveness. In the first round, there is a tie for fewest votes between *y* and *z*. If *y* is eliminated, then in the next round of counting *z* beats *x*. Since the addition of a ranking placing *x* first did not turn a perfect tie into a win for *x*, two-ballot majority lacks perfect-tie responsiveness.

Example 13.8 shows that two-ballot majority lacks combinativity. For sequence **R**, two-ballot majority eliminates *z* and then elects *x*. For sequence **R′** it elects *x* immediately. But for the sequence of rankings formed by appending **R′** to **R**, it eliminates *y* and then elects *z*. Thus two-ballot majority lacks combinativity.

Example 13.12 shows that two-ballot majority lacks independence of clones.

Example 13.12

2	3	4	5
w	*x*	*y*	*z*
x	*w*	*z*	*w*
y	*y*	*w*	*x*
z	*z*	*x*	*y*

Options *w* and *x* are clones, and under two-ballot majority there is run-off between *y* and *z*, which is won by *y*. But if *w* is removed from consideration, there is a run-off between *x* and *z*, which is won by *z*.

Turning to qualitative attractiveness, I give two-ballot majority a score of 9 for lucidity, just below plurality, and a score of 8 for ease of use, reflecting the need to vote twice. I give it a score of 8 for cost of computation, reflecting the fact that two plurality counts are needed.

The strategy requiring the fewest strategants that is sometimes successful against two-ballot majority is to give enough additional votes to the option that comes in third in the first round of voting that it displaces one of the other two options. If there is an election in which *x*, *y*, and *z* come in first, second and third respectively in the first round of voting, then the strategy involves giving *z* enough additional votes to displace *y*. If *z* then wins in the second round of voting, the strategy is a direct hoisting strategy that might be undertaken by voters who prefer z to *x* but did not vote for *z* in the first round. If *y* would win a run-off between *x* and *y*, but *x* wins a run-off between *x* and *z*, then the strategy is an indirect hoisting strategy that might be undertaken by voters who favor *x* to *y*—either voters who did not vote for *x* or some of those who did vote for *x*, if switching their votes to *z* does not cause *x* to fall into third place in the first round of voting.

For the 87 elections in the sample, there were 63 in which the number of voters who were needed to implement the minimal strategy under two-ballot majority was greater than the number available, and the rule was therefore immune to strategy. Only 24 elections permitted strategy. Figure 13.6 shows a scatter plot, for the 87 elections in the sample, of the percentages of voters available and needed for the minimal two-ballot strategy to succeed. The resistance to strategy of two-ballot majority rule is 8.1.

Overall, two-ballot majority improves on plurality in reducing the incentive to vote strategically and in providing a winning option that beats at least one other option by a majority, but it still leaves much to be desired.

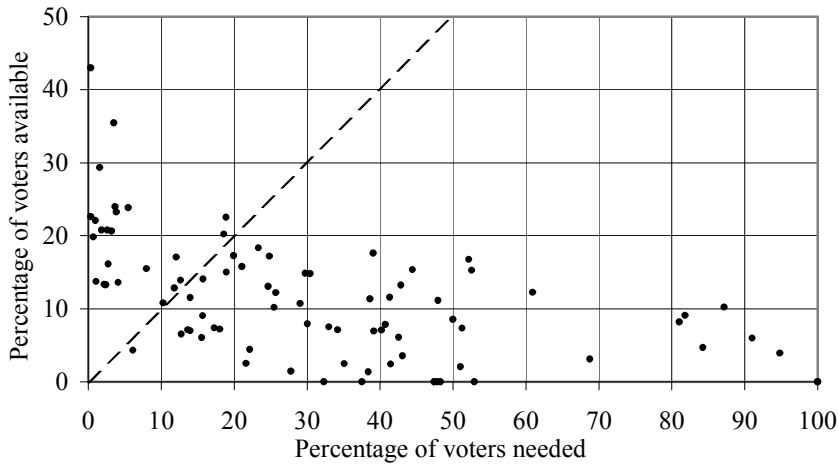


Figure 13.6 Scatter plot of the percentages of voters needed and available for minimal strategies under two-ballot majority rule, for 87 elections

The Alternative Vote

The alternative vote can be described as the application of the single transferable vote (mentioned in Chapter 9 and discussed in detail in Chapter 15), which is generally employed to elect multiple options, to the election of one option. This involves an intricate algorithm. When the rule is reduced to one for electing just one option it is quite straightforward.

The selection of a winner by the alternative vote begins by applying the plurality rule to the voters' rankings of the options. If any option receives a majority of the votes, it wins. If no option receives a majority, then the option with the fewest votes is eliminated, and its votes are allocated to the options named second on them. If an option now has a majority, that option is the winner. If not, the option with the fewest votes is eliminated, its votes are transferred to the remaining option ranked highest on them. The search for an option with a majority and the transfer of the votes of the option with the fewest votes continue until some option has a majority of the unexhausted ballots.

The alternative vote possesses the properties of mutual majority consistency (and therefore majority consistency), invariant loss consistency, homogeneity and independence of clones, but it lacks Condorcet consistency (and therefore Smith consistency and Schwartz consistency), monotonicity (and therefore positive responsiveness), resolvability, perfect-tie responsiveness, and combinativity.

It possesses mutual majority consistency (and therefore majority consistency) because if there is a set of options that are ranked ahead of all other options by a

majority of the voters, then whenever one of those options is eliminated, all of its votes are transferred to other options in the set, so that some option in the set must eventually get a majority of the votes. It possesses invariant loss consistency because an option that loses to all other options cannot win in the last round. The alternative vote possesses homogeneity because repeating a sequence of rankings k times multiplies all alternative vote scores at every stage of the count by a factor of k. The reason that the alternative vote possesses independence of clones is that if an election includes a set of clones, C, then whenever an option in C is eliminated, all of its votes are transferred to other options in C, as long as any of them remain uneliminated. Furthermore, the receipt by options in C of votes from the elimination of other options does not depend on how many options are in C. Thus when just one option in C remains uneliminated, that option will have exactly as many votes, and exactly the prospects of receiving votes from the elimination of other options, that it would have had if it had been the only option from C on the ballot from the start.

Example 13.1 shows that the alternative vote lacks Condorcet consistency (and therefore Smith consistency and Schwartz consistency), since the alternative vote and two-ballot majority are equivalent for examples with three options. Similarly, Example 13.10 shows that the alternative vote lacks monotonicity (and therefore positive responsiveness), and Example 13.11 shows that it lacks resolvability.

Example 13.7 shows that the alternative vote lacks perfect-tie responsiveness.

Example 13.8 shows that the alternative vote lacks combinativity.

Turning to qualitative attractiveness, I give the alternative vote a score of 8 for lucidity, just a notch below two-ballot majority, approval voting and the Borda rule. I give it a score of 5 for ease of use, along with all other rules that require voters to submit rankings. For cost of computation, I give it a score of 6, because if there are r options, the alternative vote can require $r - 1$ plurality counts.

Strategy against the alternative vote is like strategy against two-ballot majority, except that with the alternative vote an option can be eliminated in any stage before the last, rather than in only one stage before the last. The option that wins under a given sequence of rankings is eliminated by giving enough votes to enough other options that the winner under the given rankings becomes the option with the fewest votes at some stage of the counting and is therefore eliminated. This is generally a daunting task, so that the alternative vote is quite resistant to strategy.

The strategy to which the alternative vote might be susceptible depends on whether or not the winner is a dominant option. If the winner is dominant, then the alternative vote is only susceptible to strategy if, at a stage of the vote count when there are N options remaining, the share of the eventual winner is less than $1/N$. In that case, it may be possible to vote strategically in such a way that the option that wins in the absence of strategy has the fewest votes at that stage of the counting and is eliminated.

Among the 87 elections in the sample, there was a dominant option in 85 cases, and in 82 of these cases the dominant option was selected by the alternative vote.

In one case there was a tie between two options that both beat all others, and the alternative vote produced a tie between these two. In the final case, the alternative vote produced a tie between an option that was otherwise dominant and an option that lost to one other.

In the two cases without a dominant option, the result was not susceptible to strategy because the tied winners were never close to being the option with the fewest votes. In the three cases in which there was a dominant option that did not win, the result was strongly susceptible to strategy because a change of just a few votes would have kept the dominant option from ever being the one with the fewest votes. This is shown in Figure 13.7. The resistance to strategy of the alternative vote is 9.7.

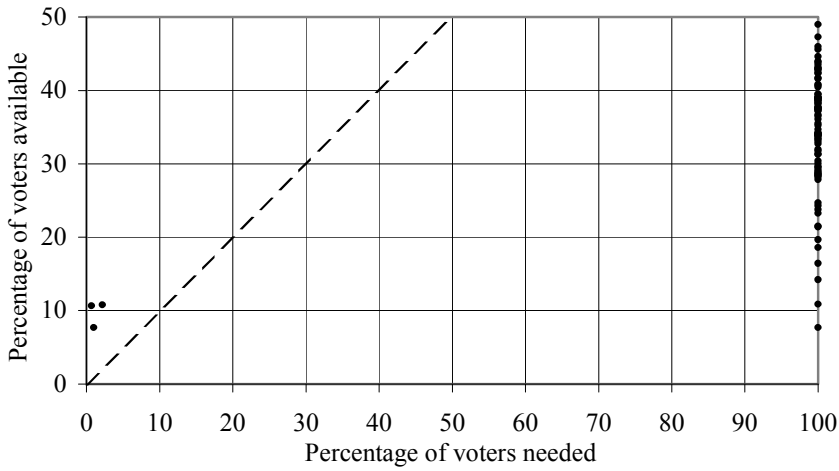


Figure 13.7 Scatter plot of the percentages of voters needed and available for minimal strategies under the alternative vote, for 87 elections

The lack of monotonicity in the alternative vote should probably not be considered a serious defect since many competing rules lack this property as well. The lack of resolvability is also probably not too serious, because it seems likely that irresolvable ties would be quite rare. The lack of Condorcet consistency is more troubling. Over the past two centuries numerous voting scholars, such as Condorcet, Dodgson, Nanson and Black, have held that Condorcet consistency is a prime requirement of a good vote-processing rule. The fact that in 87 elections in the sample there were just three in which there was a dominant option that was not chosen is somewhat reassuring.

The Dodgson Rule

Charles Dodgson, known best for writing *Alice in Wonderland* under the pen name Lewis Carroll, was a mathematician at Oxford University from the 1850s to the 1890s, who also wrote extensively on committee decisions, elections and proportional representation.[15] Dodgson used a matrix somewhat like the matrix of paired comparisons used here, except that each component was expressed as a fraction in which the numerator was the number of voters who ranked the row option ahead of the column option, and the denominator was the number of voters who ranked the column option ahead of the row option. Fractions less than 1 were placed in parentheses.[16] Like Condorcet, whose work he was unaware of, Dodgson felt that if there was a dominant option, then that option should be the winner.[17] If there was no dominant option, Dodgson asserted that there should be further discussion, and if cycles persisted, the outcome should be declared "no Election."[18] He discussed two other ways of dealing with cycles, selecting the plurality winner and selecting the winner under the alternative vote.[19] It is not clear whether these methods were meant to be applied to all options or just to the options involved in a cycle. In discussing why these alternatives to "no Election" are unattractive, Dodgson appears to employ the rule that any given option should be scored according to the number of inversions of adjoining options in individual rankings that would be needed to make the given option dominant, and the option with the lowest such score should be the winner. This has led to such a rule being called the "Dodgson rule."[20] Applying the Dodgson rule to Example 13.4, one observes first from the matrix of majorities that *v* loses to *y* and to *z* by 14 votes. The voters in the last column of the example rank *v* immediately below *y* and *z*, and if 8 of these were to revise their opinions of *v* and place it at the tops of their rankings, *v* would then be a dominant option. (Each inversion changes the majority for the inverted pair by two.) This would involve a total of 16 inversions of adjoining options. Accordingly, option *v* is given a "Dodgson score" of 16. Option *w* needs 8 voters from the third column to move *w* up by two positions, and two from either the first column or the last column to move it up one position, to become dominant. Thus *w*'s Dodgson score is 18. Option *x* needs a total of 10 voters from the first, second or fourth columns to move *x* above both *v* and *w* to become dominant. Thus *x*'s Dodgson score is 20. Option *y* needs the 9 voters from the first column to move it above *x* to become dominant. Thus *y*'s Dodgson score is 9. Option *z* needs 9 voters from the second or third columns to move it above *x* and two voters from the first or last columns to move it above *y* to become dominant. Thus *z*'s Dodgson score is 11. Option *y* with a Dodgson score of 9 is thus the winner by the Dodgson rule for this example.

The Dodgson rule possesses Condorcet consistency (and therefore majority consistency). It lacks mutual majority consistency, Smith consistency (and therefore Schwartz consistency), invariant loss consistency, resolvability, perfect-tie

responsiveness, monotonicity (and therefore positive responsiveness), homogeneity, combinativity, and independence of clones.

The Dodgson rule is Condorcet consistent because a dominant option will have a Dodgson score of 0, and since the dominant option beats every other option, every other option will have a positive Dodgson score.

Example 13.13 shows that the Dodgson rule lacks mutual majority consistency.

Example 13.13

4	2	5	3	2	4	2	2	3
w	*w*	*x*	*y*	*y*	*z*	*z*	*x*	*z*
x	*z*	*y*	*w*	*z*	*w*	*x*	*z*	*y*
y	*x*	*w*	*x*	*w*	*x*	*y*	*y*	*w*
z	*y*	*z*	*z*	*x*	*y*	*w*	*w*	*x*

The matrix of majorities for Example 13.13 is

	w	*x*	*y*	*z*
w	0	9	−7	1
x	−9	0	11	1
y	7	−11	0	1
z	−1	−1	−1	0.

Options *w*, *x*, and *y* have a mutual majority over option *z*, but the Dodgson rule selects *z* by reversing one pair on each of three ballots, while the other three options require more reversals to become dominant.

Example 13.13 also shows that the Dodgson rule lacks Smith constancy, and therefore Schwartz consistency. The Smith set for Example 13.13 is $\{w, x, y\}$, but the Dodgson rule selects *z*. Example 13.13 shows as well that the Dodgson rule lacks invariant loss consistency.

That the Dodgson rule lacks monotonicity (and therefore positive responsiveness) can be seen from Example 13.14.

Example 13.14

2	1	1	2	1	2
w	*x*	*x*	*x*	*y*	*y*
y	*w*	*w*	*z*	*w*	*z*
z	*y*	*z*	*w*	*x*	*x*
x	*z*	*y*	*y*	*z*	*w*

The matrix of majorities for Example 13.14 is

	w	x	y	z
w	0	−3	3	1
x	3	0	−1	1
y	−3	1	0	3
z	−1	−1	−3	0.

The Dodgson scores for w, x and y are all 2 because w and x are both ranked immediately below the options to which they lose at least twice, but x is never ranked immediately below y, and therefore must be raised twice in one ranking to become dominant. But consider what happens if the voter represented by the next to last column interchanges options w and y. The only change in the matrix of majorities is that $M(w, y)$ increases from 3 to 5, raising y's Dodgson score to 3. But the change puts x just below y in one ranking, thereby lowering x's Dodgson score to 1, while w's Dodgson score remains 2, so that x now wins. Since the increase in w's position changed w from a tied winner to a losing option, the Dodgson rule lacks monotonicity, and therefore positive responsiveness.

The reason that the Dodgson rule is not resolvable is that both in a case in which an option loses to another by one vote and in a case in which two options are tied, one inversion of adjoining options in a ranking is required to cause the option that loses to win. Thus in the simplest cycle of three options for three voters like Example 9.1, there is no way to make an option dominant and change its Dodgson score from 1 to 0 by the addition of one vote. The Dodgson rule is therefore not resolvable. The lack of resolvability is not as serious as it might be, because every tie can be resolved in favor of a given tied option by the addition of either one or two votes that show the given option ahead of all other options tied.

That the Dodgson rule lacks perfect-tie responsiveness can be seen from Example 13.7, which yields a tie between x and z by the Dodgson rule, both of which have Dodgson scores of 2.

That the Dodgson rule lacks homogeneity (and therefore combinativity) can be seen from Example 13.15.

Example 13.15

3	3	1	5
w	x	z	z
x	y	x	y
y	w	y	w
z	z	w	x

The matrix of majorities for Example 13.15 is

	w	*x*	*y*	*z*
w	0	4	−2	0
x	−4	0	6	0
y	2	−6	0	0
z	0	0	0	0

The Dodgson rule selects *w*, but if the rankings leading to this matrix are repeated three times, so that the components of the matrix triple, then *z* is the winner. Thus the Dodgson rule lacks homogeneity, and therefore combinativity.

Example 13.4 shows that the Dodgson rule lacks independence of clones. As was mentioned previously, the winner for this example under the Dodgson rule is *y* with a Dodgson score of 9. But if *z*, a clone of *y*, is eliminated, then the Dodgson scores for *v*, *w*, *x* and *y* become 8, 10, 20 and 9 respectively, making *v* the winner.

Turning to qualitative attractiveness, I give the Dodgson rule a score of 6 for lucidity (equal to that of the Condorcet rule) and, like all rules that require rankings, a score of 5 for ease of use. Its score of 3 for cost of computation (equal again to that of the Condorcet rule), arises from the fact that it could be necessary to examine numerous alternative combinations of inversions to determine the minimum number of inversions needed to make an option dominant.

The strategic possibilities for the Dodgson rule are like those of the Condorcet rule, involving the creation of cycles. Because the calculations for the Dodgson rule are so tedious, and because the simplified Dodgson rule, to be discussed next, is so much more attractive, I made no explicit strategy calculations for it.

The Simplified Dodgson Rule

Because the Dodgson rule depends not only on the matrix of majorities, but also on whether options that are close to being dominant are ranked immediately below the options to which they lose, its calculation can be extremely tedious. Some of the anomalies of the Dodgson rule can be eliminated at the same time that its calculation is simplified by using the following Dodgson-like rule. Define an option's simplified Dodgson score, for a given matrix of majorities, as the sum of the negative components of the option's row of the matrix of majorities. The simplified Dodgson scores of options *v*, *w*, *x*, *y* and *z* for Example 13.4 are −28, −30, −36, −16 and −18. The simplified Dodgson score is −2 times what the Dodgson score would be if:

1. The goal were undominatedness rather than dominance,
2. Every option appeared often enough immediately below the options to which it lost that it was never necessary to waste an inversion getting up to those options, and

3. The creation of a tie between adjoining options or the breaking of a tie between adjoining options counted as ½ in the Dodgson score, with an inversion counting as 1.

The simplified Dodgson rule possesses the properties of Condorcet consistency (and therefore majority consistency), resolvability, monotonicity and homogeneity, but not positive responsiveness, perfect-tie responsiveness, Smith consistency (nor therefore Schwartz consistency), independence of clones, or combinativity.

The proof that the simplified Dodgson rule is Condorcet consistent is essentially the same as the proof for the Dodgson rule.

The reason that the simplified Dodgson rule is resolvable is that if one adds a vote that ranks a given option first followed by all others tied, the only effect on the matrix of majorities is to increase by one all components of the row of the given option and decrease by one all components of the corresponding column. This increases (makes less negative) the simplified Dodgson score of the given option if it is negative and either decreases or leaves unchanged all other simplified Dodgson scores. The original Dodgson rule would have been resolvable if an inversion of the order of two adjoining options were considered a sequence of two operations, namely the creation of a tie between the two and the breaking of that tie, and each operation was assigned a "cost" of ½.

The reason that the simplified Dodgson rule possesses monotonicity is that the inversion of adjoining options i and j in one ranking increases $M(i, j)$ by 2 and decreases $M(j, i)$ by 2. Thus it can raise by up to 2 the simplified Dodgson score of option i, lower by up to 2 the simplified Dodgson score of option j, and it will have no other effects. This cannot cause a winning option to tie or lose, or cause a tied winner to lose.

The simplified Dodgson rule is homogeneous because repeating a set of rankings k times multiplies all simplified Dodgson scores by a factor of k.

The same examples that showed that the Dodgson rule lacks mutual majority consistency, Smith consistency (and therefore Schwartz consistency), invariant loss consistency, perfect-tie responsiveness, and independence of clones also show that the simplified Dodgson rule lacks these properties.

The simplified Dodgson rule lacks positive responsiveness because if two options are tied by the simplified Dodgson rule, an improvement in the score of a tied option relative to an option that it already beats and that has a lower simplified Dodgson score will not affect the tie.

Example 13.8 shows that the simplified Dodgson rule lacks combinativity. Under the simplified Dodgson rule, x wins for sequence $\mathbf{R}$ and for sequence $\mathbf{R}'$, but z wins for sequence $\mathbf{R} + \mathbf{R}'$.

In terms of qualitative attractiveness, I give the simplified Dodgson rule scores of 8 for lucidity (the same as the alternative vote), 5 for ease of use (like all rules that require rankings), and 7 for cost of computation (since it requires only a simple computation from the matrix of majorities).

As long as there are dominant options, strategic opportunities under the simplified Dodgson rule are exactly the same as under the Condorcet rule. For the 87 elections in the sample, there were just two that lacked dominant options, and in those cases the result was sufficiently close to dominant that strategic opportunities remained the same as for the Condorcet rule. Therefore the graph of strategic opportunities under the simplified Dodgson rule is equivalent to Figure 13.5, and the resistance to strategy of the simplified Dodgson rule is 8.9.

The Nanson Rule

E.J. Nanson was a Professor of Mathematics in Melbourne, Australia from 1875 to 1922.[21] Nanson accepted Condorcet consistency as a primary requirement in a vote-processing rule and proposed a combination of the Borda rule and the alternative vote to achieve it: At each stage eliminate the option or options with below-average Borda scores, calculated with reference to options that have not yet been eliminated.[22] If adjusted Borda scores (row sums of the matrix of majorities) are substituted for Borda scores, this rule is quite simple to employ, because below-average adjusted Borda scores are negative. Thus one would eliminate all options with negative row sums of the matrix of majorities, delete the corresponding rows and columns of the matrix of majorities, recompute the row sums for the remaining options on the reduced matrix, delete the options with negative row sums, and so on until either just one option remained or all options had row sums of zero (meaning there was a tie among these options).

The Nanson rule possesses mutual majority consistency (and therefore majority consistency), Smith consistency (and therefore Condorcet consistency), invariant loss consistency and homogeneity, but it lacks Schwartz consistency, monotonicity (and therefore positive responsiveness), perfect-tie responsiveness, combinativity, and independence of clones.

I do not know whether the Nanson rule is resolvable. Because the rule eliminates all options with below average Borda scores, and because the addition of one more ranking generally changes the average Borda score of options, it may not always be possible to add a ranking without changing the sequence of eliminations prior to the final tie, making it impossible to resolve some ties in favor of some options by the addition of a vote.

The Nanson rule possesses mutual majority consistency because if a set of options has a mutual majority they always all have positive components of their rows of the matrix of majorities with respect to all options outside their set. With respect to comparisons with one another, the sum of the components of the matrix of majorities is always zero. Therefore there must at all times be at least one member of the set with a mutual majority that has positive adjusted Borda score and therefore cannot be eliminated by the Nanson rule. Therefore the Nanson rule possesses mutual majority consistency and therefore majority consistency.

The Nanson rule is Smith consistent because every comparison of an option in the Smith set with one that is not produces a positive component of the matrix of majorities for the option in the Smith set, while comparisons within the Smith set produce components that average zero, so that, no matter how many options have been eliminated, there must always be at least one member of the Smith set with a positive adjusted Borda score. Thus the Nanson rule is Smith consistent, and therefore Condorcet consistent.

The Nanson rule possesses invariant loss consistency because an option that loses to all other options has a negative adjusted Borda score and will be eliminated in the first round of vote counting.

The Nanson rule is homogeneous because repeating a sequence of rankings k times multiplies all adjusted Borda scores at each stage by a factor of k.

Example 13.16 shows that the Nanson rule lacks Schwartz consistency.

Example 13.16 (Matrix of Majorities)

	v	w	x	y	z
v	0	8	−2	0	0
w	−8	0	6	0	0
x	2	−6	0	2	0
y	0	0	−2	0	4
z	0	0	0	−4	0

The GOCHA set is $\{v, w, x\}$, but the Nanson rule produces a tie between v and y.

Example 13.17 shows that the Nanson rule lacks monotonicity (and hence positive responsiveness).

Example 13.17

$\underline{1}$	$\underline{1}$	$\underline{1}$	$\underline{2}$
x	y	y	z
y	x	z	x
z	z	x	y

The matrix of majorities for Example 13.17 is

	x	y	z
x	0	1	−1
y	−1	0	1
z	1	−1	0.

Now suppose the voter represented by the second column of Example 13.17 interchanges the positions of x and z. The matrix of majorities becomes

	x	y	z
x	0	1	−3
y	−1	0	1
z	3	−1	0.

Now x is eliminated in the first stage, and in the second stage y beats z. Thus the improvement in z's position caused it to change from a tied winner to a loser. As in the case of the alternative vote, the example that demonstrates the lack of monotonicity of the Nanson rule involves a cycle. Since the Nanson rule is Condorcet consistent, any example of lack of monotonicity for this rule would involve a cycle.

Example 13.7 shows that the Nanson rule lacks perfect-tie responsiveness. The Nanson rule eliminates y in the first round and then eliminates x, although x must win for perfect-tie responsiveness.

Example 13.8 shows that the Nanson rule lacks combinativity. Under the Nanson rule, z is eliminated and then x wins for sequence **R**, while y is eliminated and then x wins for sequence **R**′, but for sequence **R** + **R**′, y is eliminated and then z wins.

Example 13.4 shows that the Nanson rule lacks independence of clones. For this example under the Nanson rule, v, w and x are eliminated in the first stage and z in the second stage, leaving y the winner. But without w, a clone of v, the options eliminated in the first round would be v and z, and in the second round x would beat y to become the winner.

In terms of qualitative attractiveness, I give the Nanson rule scores of 7 for lucidity (higher than the Condorcet rule and the Dodgson rule, lower than the alternative vote and the simplified Dodgson rule), 5 for ease of use (like all rules that require rankings), and 6 for cost of computation (equal to the alternative vote, which also involves successive eliminations).

The strategic opportunities under the Nanson rule are the same as under the Condorcet rule and the simplified Dodgson rule, as long as elections have dominant options. Therefore Figure 13.5 shows the strategic opportunities, and the Nanson rule receives a score of 8.9 for resistance to strategy.

The Bucklin System

The Bucklin or Grand Junction system of voting was first adopted in Grand Junction, Colorado, in 1909, under the leadership of James W. Bucklin.[23] In 1913, the U.S. Congress prescribed Bucklin voting for electing directors of each Federal Reserve Bank.[24] In Bucklin voting, voters rank the options, and the vote count starts like in the alternative vote, with each option receiving a score equal to the number of voters who rank that option first. If there is an option with a score that is a majority of the votes, then that option is selected and the count is finished. If

no option receives a majority of the votes, then in the next round of vote counting each option receives a score equal to the number of voters who rank the option first or second. If one option has a score that is a majority of the votes, then that option is elected. If two or three options have scores that are a majority of the votes, then the winner is the option with the greatest score. If no option has a score that is a majority of the votes, then there is another round of counting, in which each option receives a score equal to the number of voters who rank the option first, second or third, and so on until at least one option has a majority of the votes.

Bucklin voting possesses mutual majority consistency (and therefore majority consistency), resolvability, monotonicity, and homogeneity. It lacks Condorcet consistency (and therefore Smith Consistency and Schwartz consistency), invariant loss consistency, positive responsiveness, perfect-tie responsiveness, combinativity, and independence of clones.

Bucklin voting possesses mutual majority consistency because if there is a set of options that are all ranked ahead of all other options by a majority of the voters, then no option can receive a majority of the votes unless it receives at least one vote from these voters, and no option can receive any votes from these voters until after all options in the set with a mutual majority have received all of the votes, by which time an option from the set with the mutual majority would have been elected. Since majority consistency is a special case of mutual majority consistency, the Bucklin rule also possesses majority consistency.

Bucklin voting possesses resolvability because a vote that ranks only a tied option will resolve a tie in favor of that option under Bucklin voting.

Bucklin voting possesses monotonicity because if some option, x, rises above some other option, y, in the ranking of a voter, then x receives a greater count in the stage of the counting corresponding to the rank to which x rose. At that stage, y receives a lower count. This change in the count cannot cause x to lose when it would otherwise win, and it might cause x to win when it would otherwise lose. Thus Bucklin voting possesses monotonicity.

Bucklin voting possesses homogeneity because repeating a set of rankings k times multiplies all Bucklin scores by a factor of k.

Example 13.18 shows that Bucklin voting lacks Condorcet consistency (and therefore Smith consistency and Schwartz consistency).

Example 13.18

2	1	1	1
x	y	z	w
y	w	y	z
z	x	x	x
w	z	w	y

For Example 13.18, Bucklin voting selects y, but x is a dominant option.

Example 13.13 shows that Bucklin voting lacks invariant loss consistency. Bucklin voting selects *z* even though it loses to all other options in paired comparisons.

Bucklin voting lacks positive responsiveness because, in the event of a tie, an upward movement of one option in one voter's ranking in a place other than where the options attain majorities will have no effect on the outcome.

Example 13.7 shows that Bucklin voting lacks perfect-tie responsiveness. The addition of the last vote leaves a tie between *x* and *y* under Bucklin voting.

Example 13.19 shows that Bucklin voting lacks combinativity.

Example 13.19

Sequence **S**			Sequence **S′**	
6	5	4	5	4
x	*y*	*z*	*x*	*z*
z	*x*	*y*	*y*	*z*
y	*z*	*x*	*y*	*x*

Sequence **S** elects *x* on the second round of counting by the Bucklin system. Sequence **S′** elects *x* on the first round of counting by the Bucklin system. But sequence **S** + **S′** elects *y* on the second round of counting by the Bucklin system. Thus the Bucklin system lacks combinativity.

Example 13.4 shows that the Bucklin system lacks independence of clones. For Example 13.4, the Bucklin system elects z in the second round of counting. But if *v*, a clone of *w*, is deleted, the Bucklin system elects *x* in the second round of counting. Thus the Bucklin system lacks independence of clones.

With respect to qualitative criteria, I give the Bucklin system a score of 8 for lucidity, equal to that of the alternative vote and the simplified Dodgson rule. It gets a score of 5 for ease of use, like all rules that require rankings. The cost of computation is scored at 6, like the alternative vote and the Nanson rule.

Strategy in the Bucklin system is like strategy in approval voting. A strategic voter will seek to determine the two options that are most likely to win and give a vote to only one of them. The number of voters who are needed to implement a strategy is usually the difference between the vote totals of the two options in the round in which the counting ends. However, it is possible for a strategy to change the number of rounds of counting, which makes strategic calculations more complex.

When there is just one round of voting, Bucklin voting reaches the same result as plurality and offers the same strategic opportunities. When voters rank so few options that no option ever receives a majority of the votes, Bucklin voting is equivalent to approval voting and has the same strategic opportunities. Usually the strategic opportunities lie between those of these two systems. Among the 87 elections in the sample, there were 29 that were impervious to strategy under Bucklin voting. The scatter plot of voters needed and available for minimal

strategies under Bucklin voting is shown in Figure 13.8. The resistance to strategy of Bucklin voting is 5.8.

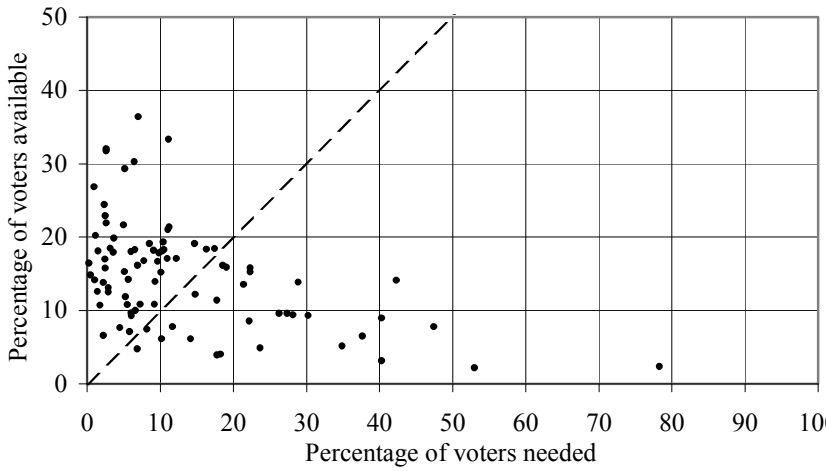


Figure 13.8 Scatter plot of the percentages of voters needed and available for minimal strategies under Bucklin voting, for 87 elections

The Copeland Rule

In a mimeographed paper in 1951, A.H. Copeland suggested that a reasonable social welfare function could be obtained by scoring options according to the difference between the number of other options that they beat and the number to which they lost, and declaring as the winner the option(s) with the greatest score.[25] An option's Copeland score is thus the difference between the number of positive components and the number of negative components of the option's row of the matrix of majorities.

The Copeland rule possesses mutual majority consistency (and therefore majority consistency), Smith consistency (and therefore Condorcet consistency) invariant loss consistency, monotonicity, and homogeneity, but it lacks Schwartz consistency, positive responsiveness, resolvability, perfect-tie responsiveness, combinativity, and independence of clones.

The Copeland rule possesses mutual majority consistency by the following argument. Suppose there are k options that are not in the set with a mutual majority, M. The option in M with the greatest Copeland score has a Copeland score of at least k, while the option not in M with the greatest Copeland score has a Copeland score of at most $k - 1$. Therefore the Copeland winner must be in M. Since the Copeland rule possesses mutual majority consistency, it possesses majority consistency.

The Copeland rule possesses Smith consistency (and therefore Condorcet consistency) because any option in the Smith set must have a higher Copeland score than any option outside the Smith set. For every option outside the set, the option inside the Smith set has 1 added to its Copeland score, where the option outside the set has at least one zero (for itself). For every option inside the Smith set, the option outside the set has – 1 added to its Copeland score, where the option inside the set has at least one zero (for itself). Thus the sum for the option inside the Smith set must exceed the sum for the option outside the Smith set.

The Copeland rule possesses invariant loss consistency because an option that loses to all other options has a Copeland score of $1 - n$, where n is the number of options, while the average Copeland score is 0, so that the option that loses to all others cannot have the greatest Copeland score.

The Copeland rule possesses monotonicity because an inversion of a pair of adjoining options in one ranking increases one component of the row of the matrix of majorities of the option that rises and decreases one component of the row of the option that loses, which *might* increase the Copeland score of the one and decrease the Copeland score of the other, and would have no other effects.

The Copeland rule is homogeneous because repeating a sequence of rankings k times has no effect on any Copeland score.

Example 13.20 shows that the Copeland rule lacks Schwartz consistency.

Example 13.20 (Matrix of Majorities)

	v	*w*	*x*	*y*	*z*
v	0	2	0	0	0
w	– 2	0	2	2	2
x	0	– 2	0	2	2
y	0	– 2	– 2	0	2
z	0	– 2	– 2	– 2	0

The GOCHA set is $\{v\}$, but the Copeland winner is w.

That the Copeland rule lacks positive responsiveness can be seen by deleting the last column from Example 13.7. Since every paired comparison is now decided by 3 votes, any interchange of the positions of two adjoining options in one ranking will leave the signs of all components of the matrix of majorities unchanged, thus leaving the three-way tie unresolved by the Copeland rule. The fact that the introduction of the last column into Example 13.7 does not affect the outcome also shows that the Copeland rule lacks resolvability and perfect-tie responsiveness.

The lack of resolvability in the Dodgson rule was not terribly serious because, although some ties could not be resolved by the addition of one vote, all ties could be resolved by the addition of two votes. The lack of resolvability in the Copeland rule is more serious. The Copeland rule calls every election involving a simple

cycle among three options a tie. The only limit on the number of additional votes that could be required to resolve a tie is that it could not require more than one-third of the number already counted. If *the* central problem of vote-processing rules is the possibility of cycles, a rule that fails to identify a unique winner whenever there is a simple cycle among three options is seriously defective.

Example 13.21 shows that the Copeland rule lacks combinativity:

Example 13.21 (Matrix of Majorities)

Matrix **M**

	x	*y*	*z*
x	0	2	−1
y	−2	0	6
z	1	−6	0

Matrix **M′**

	x	*y*	*z*
x	0	−3	5
y	3	0	−4
z	−5	4	0

Matrix **M** + **M′**:

	x	*y*	*z*
x	0	−1	4
y	1	0	2
z	−4	−2	0

The Copeland rule produces three-way ties for matrices **M** and **M′**, but a win for option *y* for matrix **M** + **M′**

Example 13.4 shows that the Copeland rule lacks independence of clones. Under the Copeland rule the winner of this election is *y*. But without the clone *w* the result would be a tie between *x* and *y*. Since the tie could be resolved in favor of *x*, the Copeland rule is not independent of clones.

In terms of qualitative attractiveness, I give the Copeland rule scores of 8 for lucidity (equal to the alternative vote and the simplified Dodgson rule), 5 for ease of use (like all rules that require rankings) and 7 for cost of computation (since the result can be determined from a simple calculation from the matrix of majorities).

With respect to strategy, the Copeland rule is somewhat like the Condorcet rule, except that it is not possible to guarantee that an option will win by creating a cycle; one can only create a 1/3 probability that each of the options in a three-way cycle will win. For this reason, some persons who favor an option for which a strategy is being developed would decide that the strategy was unacceptably risky. If *x* will win in the absence of strategy and strategy produces a 1/3 probability of a win for each of *x*, *y*, and *z*, then a person who regards *z* to be much worse than *x* or *y* may decline to participate in the strategy even if he or she prefers *y* to *x*. Therefore I counted as strategants only person who preferred both of the options that the strategy added as potential winners.

Figure 13.9 is a scatter plot of the percentages of voters needed and available to implement minimal strategies under the Copeland rule.

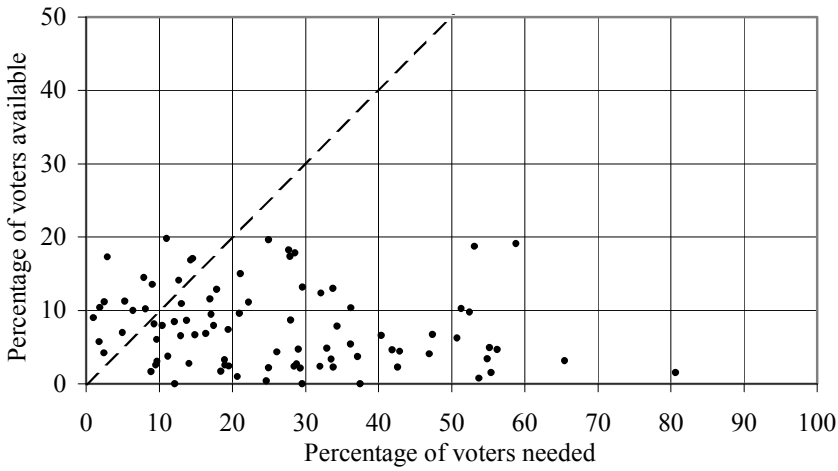


Figure 13.9 Scatter plot of the percentages of voters needed and available for minimal strategies under the Copeland rule, for 87 elections

The resistance score for the Copeland rule calculated in terms of the expected proportion of voters who must be contacted to achieve an effective coalition is 9.1. However, an effective coalition produces a 1/3 probability of success for each of three options, two of which are favored by the coalition. On this basis, the score of the Copeland rule for resistance to strategy is increased by 1/3 of the distance to 10, from 9.1 to 9.4.

The Black Rule

Duncan Black felt that when there was a dominant option, that option should win.[26] When there was no dominant option, Black advocated a rule that the option with highest adjusted Borda score is the winner.[27] Thus the Black rule is the two-part rule that when there is a dominant option it wins, and otherwise adjusted Borda scores are used to rank options.

The Black rule possesses the properties of Condorcet consistency (and therefore majority consistency), invariant loss consistency, positive responsiveness (and therefore monotonicity), resolvability, perfect-tie responsiveness, and homogeneity. It lacks mutual majority consistency, Smith consistency (and therefore Schwartz consistency), combinativity, and independence of clones.

The Black rule possesses Condorcet consistency (and therefore majority consistency) by its definition. It possesses invariant loss consistency because if there is a dominant option it wins, and if not, an option with invariant losses will

have a negative adjusted Borda score and therefore cannot have the greatest Borda score, since the average adjusted Borda score is zero.

The Black rule is positively responsive (and hence monotonic) for the following reasoning: If there is a winner by dominance, an incremental upward movement by the dominant option will leave it dominant, and if there is a winner by adjusted Borda scores, an incremental upward movement by the winning option will improve its score and reduce one other score, leaving it the winner. The existence of a tie would mean that there was no dominant option and two or more options had the same adjusted Borda score. An incremental upward movement for any given tied option would increase its adjusted Borda score and decrease one other adjusted Borda score. It could not possibly make any other option dominant, so it would resolve the tie in favor of the given option. Thus the Black rule is positively responsive, and therefore monotonic.

The Black rule is resolvable because if there is a tie, the addition of a ranking with one of the tied options first, followed by all other options tied in second place, will resolve the tie in adjusted Borda scores in favor of the option ranked first, and it cannot make any other option dominant.

The Black rule is perfect-tie responsive because, in a perfect tie, all options have adjusted Borda scores of 0. Therefore the augmentation of a perfect tie by a ranking that places one option first will give that option the greatest adjusted Borda score and make it the winner.

The Black rule is homogeneous because repeating a sequence of rankings k times leaves dominance unaffected and multiplies all adjusted Borda scores by a factor of k.

That the Black rule lacks mutual majority consistency is shown by Example 13.9, where x, y, and z have a mutual majority, but, since no option is dominant, w wins with a Borda score of 51.

Example 13.22 shows that the Black rule lacks Smith consistency (and therefore Schwartz consistency).

Example 13.22 (Matrix of Majorities)

	v	w	x	y	z
v	0	3	−3	1	1
w	−3	0	3	1	1
x	3	−3	0	1	1
y	−1	−1	−1	0	7
z	−1	−1	−1	−7	0

The Smith set for Example 13.22 is $\{v, w, x\}$, but the Black winner is y.

Example 13.8 shows that the Black rule lacks combinativity. Under the Black rule, x wins for sequence **R** because it has the greatest adjusted Borda score, and x wins by dominance for sequence **R′**, but z wins by dominance for sequence **R** + **R′**.

Since there is no dominant option in Example 13.4, the Black rule yields the same result for this example as the Borda rule. Thus the proof that the Borda rule lacks independence of clones also shows that the Black rule lacks independence of clones.

In terms of qualitative attractiveness, I give the Black rule scores of 7 for lucidity (equal to the Condorcet rule and the Nanson rule), 5 for ease of use (like all rules that require rankings) and 7 for cost of computation (since the winner can be determined from a simple computation from the matrix of majorities).

The strategy that might be employed by supporters of the option ranked second when the Black rule is in use would be an indirect hoist of the option ranked third to create a cycle, combined with a direct hoist of the option ranked second to give it enough extra Borda points to cause it to win. Figure 13.10 shows a scatter plot of the percentages of voters needed and available to succeed with such a strategy in the 87 elections in the sample. The resistance to strategy of the Black rule is 7.8.

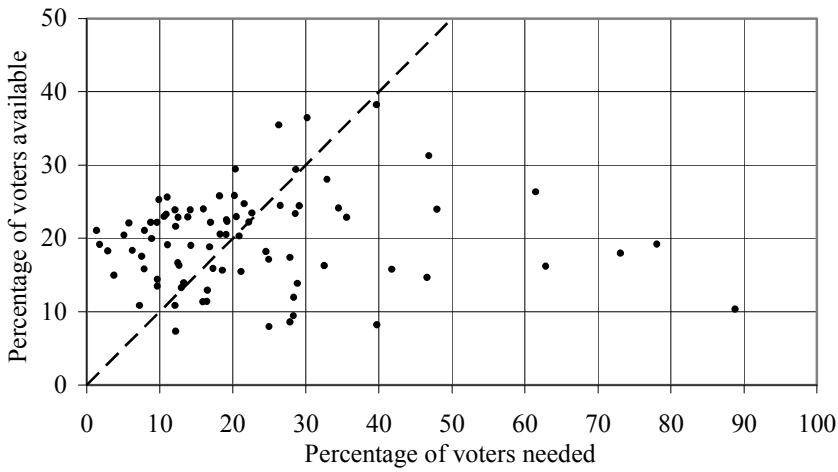


Figure 13.9 Scatter plot of the percentages of voters needed and available for minimal strategies under the Black rule, for 87 elections

While the Black rule does reasonably well in terms of the properties under consideration, it exhibits a lack of a quality that might be called "coherence" when it says, in effect, "Do one thing, and if that doesn't work, then do something entirely different." In addition to being aesthetically unattractive, this suggests that when the rule is interpreted as an allocation of election outcome space to options, there will be not merely some non-convexity (which was shown in Chapter 12 to be true of all rules that possess Condorcet consistency and perfect-tie responsiveness), but a pronounced raggedness of the boundaries of the regions of election outcome space assigned to different options.

The Maximin Rule

In Black's effort to decipher Condorcet's proposed rule, one of the suggestions he made (based on Condorcet's erroneous algorithm for finding the winner under his rule), was that Condorcet might have intended that, in the event that there was no dominant option, the winner would be the option for which the worst loss to another option in paired comparisons was least bad.[28] I call this rule the maximin rule, based on the fact that the winner is the option whose minimum performance is better than any other minimal performance. It has sometimes been called the minimax rule, based perhaps on the fact that the rule can also be described as choosing the option whose maximum loss is smaller than any other maximum loss.

To apply the maximin rule, one scores each option according to the least (most negative) component of the option's row of the matrix of majorities (giving an option with no losses a score of zero), and the option with the greatest score (closest to 0) is the winner.

The maximin rule possesses the properties of Condorcet consistency (and therefore majority consistency), resolvability, monotonicity, and homogeneity. It lacks mutual majority consistency, Smith consistency (and therefore Schwartz consistency), invariant loss consistency, positive responsiveness, perfect-tie responsiveness, combinativity and independence of clones.

The maximin rule is Condorcet consistent because a dominant option will have a maximin score of 0 and will be the only option with a non-negative score. It is therefore also majority consistent.

It is resolvable because the introduction of a vote with a given option ranked uniquely first will increase every component of the given option's row of the matrix of majorities by 1, thereby increasing the option's maximin score if it is negative, at the same time that it decreases the option's column of the matrix of majorities, thereby decreasing or leaving unchanged every other option's maximin score.

It is monotonic because an incremental upward movement by an option can only improve the maximin score of that option and diminish the maximin scores of other options.

It is homogeneous because repeating a sequence of rankings k times multiplies all maximin scores by a factor of k.

Example 13.13 shows that the maximin rule lacks mutual majority consistency. In this example, options w, x, and y have a mutual majority over option z, but the maximin rule selects z. Since Smith set for Example 13.13 is $\{w, x, y\}$, this example also shows that the maximin rule lacks Smith consistency (and therefore Schwartz consistency). Since option z loses to all other options in paired comparisons, this example shows as well that the maximin rule lacks invariant loss consistency.

That the maximin rule lacks positive responsiveness is shown by Example 13.6, which the maximin rule calls a tie both before and after the vote switch.

That the maximin rule lacks perfect-tie responsiveness can be seen from Example 13.7, which yields a tie between x and z by the maximin rule, both of which have maximin scores of –2.

Example 13.8 shows that the maximin rule lacks combinativity. Under the maximin rule, x wins for sequence **R** and for sequence **R′**, but z wins for sequence **R** + **R′**.

Example 13.13 shows that the maximin rule is not independent of clones, since w, x and y are clones, and the elimination of w changes the winner from z to x. It is worth noting here that the maximin rule is independent of clones that come only in pairs. The introduction of a clone never changes the worst loss of a non-clone, and it has no effect on the worst loss of the one of a pair of clones that beats the other in paired comparisons.

In terms of qualitative attractiveness, I give the maximin rule scores of 7 for lucidity (equal to the Black rule and the Nanson rule), 5 for ease of use (like all rules that require rankings), and 7 for cost of computation (since the winner can be determined from a simple calculation from the matrix of majorities)

As long as there are dominant options, strategic opportunities under the maximin rule are exactly the same as under the Condorcet rule. Therefore the graph of strategic opportunities under the maximin rule is equivalent to Figure 13.5, and the resistance to strategy of the maximin rule is 8.9.

The maximin rule is a fairly good rule despite its limitations. Situations of the kind that were used to show that it lacked Smith consistency and independence of clones, where three or more options are involved in a cycle that is deeper than the margins by which all options in a cycle beat some other option, are likely to be extremely rare.

The Coombs Rule

The Coombs rule is closely related to the idea of "exhaustive voting." Duncan Black describes exhaustive voting as the following system:[29] Each voter is told to give a vote to all options but one, and the option with the fewest votes is eliminated. Then another round of voting is held in which each voter is to give one vote to every remaining option but one, and again the option with the fewest votes is eliminated, and so on until just one option remains. Black noted that when preferences are single-peaked in one dimension, every voter's last-place option (the one that does not get a vote from the voter) will be the option at one end or the other of the scale on which the options are ordered.[30] A dominant option cannot be the last choice of a majority of the voters at any stage and therefore must win. On the other hand, Black noted, if preferences are not single peaked in one dimension, as in Example 13.23, then exhaustive voting can fail to select a dominant option.

Example 13.23

3	2	2
x	*z*	*z*
y	*y*	*x*
z	*x*	*y*

Option *z* is first in a majority of the rankings in Example 13.23, but *z* also has the most last-place votes and would be eliminated first under exhaustive voting.

Clyde Coombs created "unfolding theory" to describe the way that the spatial locations of objects over which individuals have preferences could be inferred from patterns in the rankings of individuals.[31] Coombs cites some of Black's earlier work, in which he developed the idea of single-peaked preferences,[32] but Black's book, *The Theory of Committees and Elections*, is not in his bibliography. Coombs replicates Black's finding that when preferences are single peaked in one dimension, every individual will rank as last the option at one end or the other of the scale.[33] Coombs suggests that collective decisions might be made by a process in which each voter reports a ranking of the options, and in successive stages the option with the most last-place votes is eliminated.[34] Coombs notes that such a system will always select the dominant option when preferences are single-peaked in one dimension.[35]

The one way in which Coombs' suggestion differs from Black's description of exhaustive voting is that Coombs describes an RBVPR, while Black describes a voting procedure in which, at each stage, each voter would decide at that time which of the remaining options would not receive a vote from him. If voters never voted strategically, the two systems would yield precisely the same results.

The Coombs rule possesses the properties of resolvability, homogeneity, and invariant loss consistency, but it lacks majority consistency (and therefore mutual majority consistency, Condorcet consistency, Smith consistency, and Schwartz consistency), monotonicity (and therefore positive responsiveness), perfect-tie responsiveness, combinativity, and independence of clones.

The reason that the Coombs rule possesses invariant loss consistency is that if there is an option that loses to all other options in paired comparisons, then if it happens that that option is one of two left at the end of a Coombs count, it must lose in the last round.

That the Coombs rule is resolvable may be shown as follows: Suppose that one or more ties are encountered in the application of the Coombs rule to a sequence of rankings, and that for one way of resolving the tie or ties, *i* is the winner. Add a vote that ranks last the option that is eliminated first for that way of resolving the tie or ties, next to last the option that is eliminated next, and so on up to *i* first. The addition of this vote will result in the same order of elimination, but with no ties. Thus the Coombs rule is resolvable.

The Coombs rule is homogeneous because repeating a sequence of rankings k times multiplies all scores at each stage of the counting by a factor of k.

Black's example (Example 13.23) shows that the Coombs rule lacks majority consistency and therefore lacks mutual majority consistency, Condorcet consistency, Smith consistency, and Schwartz consistency.

Example 13.23 can be used to show that the Coombs rule lacks monotonicity (and therefore positive responsiveness). With the existing example, z is eliminated and then y, making x the winner. But if the two voters whose rankings are given in the second column advance x ahead of y, then y is eliminated in the first stage and z wins the election.

Example 13.7 shows that the Coombs rule lacks perfect-tie responsiveness. In the first round y is eliminated; z is then elected.

Example 13.8 shows that the Coombs rule lacks combinativity. For sequence **R**, the Coombs rule eliminates z and then elects x. For sequence **R′**, it elects x immediately. But for the sequence of rankings formed by appending **R′** to **R**, there is a tie between x and y for the most last-place votes. Whichever one is eliminated, the rule then elects z. Thus the Coombs rule lacks combinativity.

Example 13.4 shows that the Coombs rule lacks independence of clones. Under the Coombs rule the order in which the options are eliminated is x, w, v, z, with y therefore winning. But if z, a clone of y, were not present, y would be the first option eliminated.

In terms of qualitative attractiveness, the Coombs rule is given a score of 7 for lucidity (equal to the Nanson rule, the Black rule and the maximin rule), 5 for ease of use (like all rules that require rankings), and 6 for cost of computation (equal to the alternative vote and the Nanson rule, since all involve successive eliminations).

The Coombs rule offers a strong incentive to bury a disfavored contending option in last place, to try to have it eliminated early in the process of sequential eliminations, so it is likely to have a very low resistance to strategy, on the order of that of the Borda rule or approval voting. I have not computed its resistance to strategy because it did not seem worth the effort for such an unpromising rule.

The strongest claim of the Coombs rule is that it selects the dominant option when preferences are single-peaked in one dimension. However, there are other vote-processing rules that select the dominant option not just when preferences are single-peaked in one dimension but whenever there is a dominant option. Thus the claim of the Coombs rule is very weak.

The Estimated Centrality Rule

In 1976, I.J. Good and I published a description of a ranking based vote-processing rule based on a spatial model of voting.[36] We assumed that each option had a position in "attribute space" that was known to voters. The voters were assumed to have "ideal points" (points in attribute space that they would most like an option

to occupy) that were distributed according to a multivariate normal distribution. We showed that it would be possible to use the fractions of voters who placed the options in each of the possible orders to estimate which option was closest to the mode of the voters' ideal points, using a criterion of minimizing the sum of the squares of the differences between actual and predicted vote shares. We produced a computer program to do this for the case of three options. This rule will be called the estimated centrality rule.

The relationship between the voters' rankings and the estimate of which option is closest to the mode of the voters' ideal points is impractical without the use of a computer and significant programming effort, making it quite difficult to determine what formal properties the estimated centrality rule possesses. Therefore this discussion will include conjectures about its properties as well as confirmed results.

The estimated centrality rule is definitely homogeneous, and it probably possesses resolvability. It lacks Condorcet consistency (and therefore mutual majority consistency, Smith consistency, and Schwartz consistency), and combinativity, and it is conjectured to lack independent of clones. Whether it possesses majority consistency, invariant loss consistency, monotonicity and perfect-tie responsiveness remains uncertain and unconjectured. If it possesses monotonicity, it almost certainly possesses positive responsiveness.

The estimated centrality rule is definitely homogeneous because the outcome is a function of the *shares* of votes that are of different types. The reason that the estimated centrality rule probably possesses resolvability is that the relative positions that it estimates vary continuously. Therefore they are probably sensitive to any change in the relative positions of options in any ranking.

In the course of investigating the consequences of the rule for a variety of situations, I came across instances in which combinativity was not satisfied and instances in which a dominant option did not win. On this basis I can assert that the rule of estimated centrality lacks combinativity and Condorcet consistency, although the details of these examples have since been discarded. The lack of Condorcet consistency implies that it lacks mutual majority consistency, Smith consistency, and Schwartz consistency too.

The estimated centrality rule is probably not independent of clones, because the replication of an option would double the weight attached to deviations from their predicted values of vote shares of rankings that included the option.

There is little basis for speculation about whether the estimated centrality rule possesses majority consistency, invariant loss consistency, monotonicity or perfect-tie responsiveness. However, if it possesses monotonicity, then it would be reasonable to expect it to possess positive responsiveness as well, since the relative positions that it estimates vary continuously.

Turning to qualitative attractiveness, I give the estimated centrality rule a score of 3 for lucidity, half of the lowest score previously encountered. Like all rules

that require rankings, it is given a score of 5 for ease of use. The necessity of using computers in an iterated search leads to a score of 1 for cost of computation.

With respect to resistance to strategizing, the estimated centrality rule has a limitation like that of the Borda rule. By being sensitive to all of the data, the rule provides a voter with an opportunity to multiply the effect of a vote on the relative standing of two strong contenders by burying the less favored option at the bottom of a ranking and hoisting the more favored to the top. Thus the estimated centrality rule probably scores between 4 and 5 for resistance to strategizing.

The estimated centrality rule might be attractive if the necessary computing resources could be brought to bear on the problem, and if there were little or no concern for the possibility that voters would engage in strategic behavior, but not otherwise.

The Young Rule

Peyton Young suggested that the criterion that Condorcet had employed might be modified by scoring each option according to the size of the largest subset of voters' rankings for which that option was undominated, and declaring the option with the highest score the winner.[37] To give this rule homogeneity, Young specified that it would operate on fractions of rankings.

The Young rule possesses Condorcet consistency (and therefore majority consistency), resolvability, monotonicity, and homogeneity. It lacks mutual majority consistency, Smith consistency (and therefore Schwartz consistency), invariant loss consistency, positive responsiveness, perfect-tie responsiveness, combinativity, and independence of clones.

The Young rule possesses Condorcet consistency (and therefore majority consistency) because an undominated option beats or ties all other options for the entire sequence of rankings, and any other option beats or ties all others only for a proper subsequence of the rankings.

The Young rule possesses resolvability because, if there is a tie between two options, i and j, then the addition of a ranking that places i first and j last will increase i's Young score by at least 1 while adding nothing to j's Young score, thereby resolving the tie in favor of i.

The Young rule possesses monotonicity because an incremental upward movement by an option can only increase its Young score and can only decrease the Young score of the option that it passes.

The Young rule is homogeneous because, if an option has a Young score of S when a sequence of rankings is replicated k times, then for a single instance of the sequence of rankings the option can achieve a score of S/k (and can do no better) by using $1/k^{th}$ of the rankings of each type that were used when there were k replications.

Example 13.13 shows that the Young rule lacks mutual majority consistency. For example 13.13, the Young score of option z is 26, since the deletion of one voter from the first column creates a set of voters for which z is undominated. The other three options have much lower Young scores, so z is the winner by the Young rule. But w, x, and y have a mutual majority over z, so the Young rule lacks mutual majority consistency.

Example 13.13 also shows that the Young rule lacks Smith consistency (and therefore Schwartz consistency), since the Smith set for Example 13.13 is $\{w, x, y\}$, but the Young rule selects z.

Example 13.13 shows as well that the Young rule lacks invariant loss consistency, since z is the Young winner despite losing to w, x, and y in paired comparisons.

That the Young rule lacks positive responsiveness can be seen from Example 13.6, in which x and y both have Young scores of 4, and continue to both have Young scores of 4 after the voter in the third column moves x ahead of z.

Example 13.7 shows that the Young rule lacks perfect-tie responsiveness. A voting rule that possessed perfect-tie responsiveness would select x for this example, but by the Young rule it is a tie between x and z, since both of these can be made undominated by the removal of one ranking that places that option last.

Example 13.8 shows that the Young rule lacks combinativity. By the Young rule, x wins for sequence $\mathbf{R}$ and for sequence $\mathbf{R}'$, but z wins for sequence $\mathbf{R} + \mathbf{R}'$.

Example 13.13 can be used to show that the Young rule lacks independence of clones. Option w, x, and y are clones, but when w is removed, the winner under the Young rule changes from z to x.

Turning to qualitative attractiveness, I give the Young rule a score of 5 for lucidity, just below the Condorcet and Dodgson rules, and 5 for ease of use, like all rules that use rankings. For cost of computation, I give it a score of 3, equal to the Condorcet and Dodgson rules, because, like those rules, it can require the exploration of a very large number of possibilities (alternative sets of rankings, in the case of the Young rule) to determine the winner.

Since the Young rule is Condorcet consistent, it is likely to have a resistance to strategy of about 9, but I have not estimated it.

The idea behind the Young rule is an interesting one. In practice, the Young rule seems unlikely to depart often from the result of the much simpler maximin rule. The two rules will necessarily yield the same result if the number of voters who rank each option last is at least as great as the worst margin by which that option loses to any other in paired comparisons by majority rule. (The Young score for each option will be the number of voters minus the absolute value of its maximin score.) If, as seems likely, this condition is generally met, it will probably not be worthwhile to maintain the Young rule for the infrequent cases when the rules yield different results.

The Ranked Pairs Rule

Of the rules considered so far, the only one that possesses independence of clones is the alternative vote, and it lacks some important properties that most other rules possess. The ranked pairs rule was designed to achieve independence of clones while retaining other attractive properties of vote-processing rules.[38]

The idea of the ranked pairs rule is to regard the majorities in the matrix of majorities as a sequence of constraints or criteria of diminishing significance. The ranked pairs rule presumes that the most important criterion of the final result is that the complete order of the options be consistent with the largest majority between a pair of options. All orders of the options that do not preserve the majority order of this pair are eliminated. Next, all orders of the options that do not preserve the second largest majority are eliminated. When one reaches the third largest majority, it is possible that, as in a simple cycle among three options, there is no remaining order that is consistent with this criterion. In that event one has no choice but to skip this criterion and go on to the next. The process terminates when either only majorities of 0 remain, in which case the result is a tie among the options that are the top of one or more remaining orders, or only one order remains, in which case the option at the top of that order is the winner.

Consider how the ranked pairs rule would be applied to the matrix of majorities for Example 13.13, in which *w*, *x* and, *y* are clones and form the GOCHA set, but *z* wins by several rules. The greatest majority in this example is $M(x, y)$ which is 11. Therefore the selected order of the options must place *x* ahead of *y*. The second greatest majority is $M(w, x)$, which is 9. Therefore the selected order of the options must place *w* ahead of *x*. The third greatest majority is $M(y, w)$, which is 7. But it is not possible to have *y* ahead of *w* if *w* must be ahead of *x* and *x* ahead of *y*. Therefore this majority is ignored. Next, there is a three-way tie among $M(w, z)$, $M(x, z)$ and $M(y, z)$, all of which are 1. This poses no difficulty since it means simply that *z* must be placed below all three of *w*, *x* and *y*. Thus the final order by the ranked pairs rule is *wxyz*, and the winner is *w*.

To achieve independence of clones in all cases of ties, the ranked pairs rule must incorporate a specific tie-breaking procedure.[39] Ties among the majorities between pairs must be broken by a "tie-breaking ranking of the pairs," or TBRP, that is based on a "tie-breaking ranking of the options," or TBRO. Each pair is scored first by the option, between the two that compose that pair, that is ranked higher by the TBRO. When two pairs differ in terms of their option that is ranked higher by the TBRO, the TBRP ranks higher the pair whose higher-ranked option is ranked higher by the TBRO. When two pairs have the same option that is ranked higher by the TBRO, the TBRP ranks higher the pair whose other option is ranked higher by the TBRO. If two or more majorities are the same, the one that is considered first is the one between the pair with the highest ranking by the TBRP. It is as if the TBRP puts the pairs in alphabetical order, in terms of an alphabet that is specified by the TBRO. The TBRO is also used to resolve ties in paired

comparisons. The option that is higher in the TBRO is considered to have beaten the other.

To avoid rewarding cloning, the TBRO is established using a device that was discussed in Chapter 12: Pick a voter at random and use his or her ranking as the TBRO. If this ranking contains ties, pick another voter at random to resolve these, and so on until the only tied options are those that are ranked as tied by all voters. These ties are resolved by a random process.

Basing the TBRP on a TBRO serves to guarantee that if the comparison of some clone, x, with another option, i, is considered before the comparison of x with some other option, j, then the comparison of y, a clone of x, with i will be considered before the comparison of y with j. Failure to satisfy this condition can lead to instances in which independence of clones is not satisfied.[40]

When a TBRO has been selected, the ranked pairs rule generates a unique outcome, x. Any alternative outcome, y, would involve ordering y ahead of x. But this is not possible because after the algorithm considers the pair $\{x, y\}$, the unexcluded orderings must consist either entirely of orders with x ahead of y or else entirely of orders with y ahead of x.

For proving several properties of the ranked pairs rule, it is useful to employ an alternative definition of it:[41]

Let **M** be a matrix of majorities and T be a TBRO. Define a binary relation U on ordered pairs of ordered pairs by $(w, x)\ U\ (y, z)$ if and only if either $M(w, x) > M(y, z)$ or [$M(w, x) = M(y, z)$ and $\{w, x\}$ is ahead of $\{y, z\}$ in the TBRP generated by T]. Let P be a complete order of the options. Define option x to "attain" option y through order P and TBRO T if and only if there is a sequence of distinct options, $a_1, a_2, \cdots, a_j$, with $a_1 = x$ and $a_j = y$, such that $a_i\ P\ a_{i+1}$ and $(a_i, a_{i+1})\ U\ (a_j, a_1)$ for $i = 1, \cdots, j - 1$. Define a complete order of the options, P, to be a "stack" for T if and only if $x\ P\ y$ implies x attains y through P and T. An option wins according to the alternative definition of the ranked pairs rule if and only if it is ranked first in a stack for some TBRO.

If P is the complete order of options that is generated by the algorithmic definition of the ranked pairs rule using TBRO T, then P is a stack for T, and $x\ P\ y$ implies that x attains y through P and T. This is because, at the first point in the algorithm at which all complete orders that place y ahead of x have been eliminated, there must be a sequence of distinct options, $a_1, a_2, \cdots, a_j$, with $a_1 = x$ and $a_j = y$, such that $a_i\ P\ a_{i+1}$ for $i = 1, \cdots, j - 1$. If this were not the case, then one or more of the orders that placed y ahead of x would have been retained. Thus the complete order that is selected by the algorithmic definition of the ranked pairs rule for any TBRO T is a stack for T, and an option wins under the algorithmic definition of the ranked pairs rule only if it wins under the definition in terms of stacks. Thus the two definitions are equivalent.

The ranked pairs rule possesses mutual majority consistency (and therefore majority consistency), Smith consistency (and therefore Condorcet consistency), invariant loss consistency, resolvability, monotonicity, homogeneity, and

independence of clones. It lacks Schwartz consistency, positive responsiveness, perfect-tie responsiveness, and combinativity.

That the ranked pairs rule possesses mutual majority consistency follows from the definition of the rule in terms of stacks. Let x be an option in the set that has a mutual majority, while y is not in that set. If y were among the winning options, then there would be a stack that placed y first and x lower down. In such a stack there must be a pair (a_i, a_{i+1}) where a_i, is not in the set with a mutual majority and a_{i+1} is in that set. But since, in paired comparisons, every option in the set with a mutual majority beats every option not in that set, and every option in a stack beats the option that succeeds it, this cannot happen. Thus when there is a set with a mutual majority, the winner by the ranked pairs rule is an element of that set. Thus the ranked pairs rule possesses mutual majority consistency and therefore majority consistency.

A similar proof shows that the ranked pairs rule possesses Smith consistency. Let x be an option that is in the Smith set, while y is not. If y were among the winning options, then there would be a stack that placed y first and x lower down. In such a stack there must be a pair (a_i, a_{i+1}) where a_i, is not in the Smith set and a_{i+1} is in the Smith set. But since every option in the Smith set beats every option not in the Smith set in paired comparisons, and every option in a stack beats the option that succeeds it, this cannot happen. Thus the winner by the ranked pairs rule is an element of the Smith set, and the ranked pairs rule possesses Smith consistency and therefore Condorcet consistency.

The ranked pairs rule possesses invariant loss consistency because an option that loses to all other options cannot be at the top of a stack, because no other option could be second.

The proof that the ranked pairs rule is monotonic is as follows: Let $f(S, \mathbf{R}, T)$ be the function by which the ranked pairs rule picks a unique order of the options in S, given the sequence of voters' rankings $\mathbf{R}$ and TBRO T. Suppose that $f(S, \mathbf{R}, T) = Q$, that x is first in Q, and that $\mathbf{R}'$ is an incremental upward movement for x relative to $\mathbf{R}$. Let $f(S, \mathbf{R}', T) = Q'$. Let L and L' be lists of the pairs in S, ordered by the majorities by which one member of the pair beats the other for sequences $\mathbf{R}$ and $\mathbf{R}'$ respectively, with ties in the majorities broken by the TBRO T. The only way in which L' can differ from L is that one pair that includes x, say $\{x, y\}$, may have moved. If x beats y by majority rule in $\mathbf{R}$, then $\{x, y\}$ can be higher in L' than in L. If x loses to y by majority rule in $\mathbf{R}$, then $\{x, y\}$ can be lower in L' than in L. In the former case, the earlier consideration of $\{x, y\}$ would lead to the pair being placed in the order x ahead of y in Q', and pairs between the original and subsequent placements of $\{x, y\}$ would be ordered in the same way in Q' as in Q, since these pair orderings and those that preceded them are not inconsistent with x ahead of y. In the latter case, the delayed consideration of $\{x, y\}$ would lead to the pair continuing to be placed in the order x ahead of y, since that placement is necessitated by the orderings of pairs that preceded the earlier location of $\{x, y\}$, and pairs between the original and subsequent placements would be ordered the

same way in Q' as in Q, because the placement of x ahead of y is implied by the time one reaches the original location of the pair. Thus $Q' = Q$, and the ranked pairs rule is monotonic.

The proof that the ranked pairs rule is resolvable is as follows: Define f as above and let $g(S, \mathbf{R})$ be the function that identifies the set of options from set S that are chosen by the ranked pairs rule when the voters' ranking are $\mathbf{R}$. Suppose that $f(S, \mathbf{R}, T) = Q$ and that x is first in Q. Form sequence of rankings $\mathbf{R}'$ by appending Q to $\mathbf{R}$. Let Q' be any order of the options other than Q. Let M be the greatest majority, for the sequence of rankings $\mathbf{R}$, among pairs that are ordered differently in Q and Q'. Let V be the set of ordered pairs (i, j) such that i has a majority of M over j and i is ahead of j in Q and not in Q'. V must be non-empty because otherwise Q' would have been chosen rather than Q. For sequence $\mathbf{R}'$, every ordered pair in V has a majority of $M + 1$. Furthermore, there is no ordered pair (i, j) such that i has a majority of $M + 1$ over j and i is ordered ahead of j in Q' but not in Q. Therefore for every T', $f(S, \mathbf{R}', T')$ cannot be Q' or any other order of the options, so that $g(S, \mathbf{R}') = x$, and the ranked pairs rule is resolvable.

The proof that the ranked pairs rule is independent of clones involves showing that for a given TBRO, T, if one or more clones of a given option are added, they must be ordered immediately beside the option of which they are clones, and that the relative ordering of a clone and a non-clone or of two non-clones cannot be affected.

Let S be a set of options, $\mathbf{R}$ be a sequence of rankings of S, and T be a TBRO for S. Define f as above and let $D = f(S, \mathbf{R}, T)$. Let z be the first option in D.

Let C, a subset of S, be clones. Let $\mathbf{R}'$ be sequence $\mathbf{R}$ with all options not in C eliminated. Let T' be the order T with all options not in C eliminated. Let $D' = f(C, \mathbf{R}', T')$, with x the first option in D'. Let B be a set of options consisting of x plus the elements of S that are not in C. Let $\mathbf{R}''$ be $\mathbf{R}$ with all options not in B eliminated. Let T'' be order T with all options not in B eliminated. Let $D'' = f(B, \mathbf{R}'', T'')$, with y the first option in D''.

Define P to be an order of S that places the elements of C in their order in D' and all other elements of S in their order in D, with all elements of C inserted together where x appears in D''. Checking all four of the cases of ordered pairs of elements of S, where the two components of the pair are and are not elements of C, it can be seen that P is a stack for the TBRO T. Therefore $P = D$. Thus if z is not in C, then $y = z$, and if z is in C, then y is in C. Thus the ranked pairs rule is independent of clones.

That the ranked pairs rule is homogeneous follows from the fact that if all rankings are replicated k times, then every component of the matrix of majorities is multiplied by k. Therefore the ranking of pairs by majority is preserved.

Example 13.5 shows that the ranked pairs rule lacks Schwartz consistency. For this example, options w and x are tied winners under the ranked pairs rule, but w is the only element of the GOCHA set.

That the ranked pairs rule lacks positive responsiveness is shown by Example 13.6, which is a tie by the ranked pairs rule both before and after the vote switch.

That the ranked pairs rule lacks perfect-tie responsiveness can be seen from Example 13.7, which yields a tie between x and z by the ranked pairs rule, which can be broken in favor of either one, depending on the TBRO. If x is ahead of y in the TBRO, then z is the winner. If y is ahead of x, then x is the winner.

That the ranked pairs rule lacks combinativity can be seen from Example 13.8. Under the ranked pairs rule, x wins for sequence **R** and for sequence **R′**, but z wins for sequence **R** + **R′**.

Turning to qualitative attractiveness, I give the ranked pairs rule a score of 5 for lucidity, equal to the Young rule. Like all rules that require rankings, I give it a score of 5 for ease of use. I give it a score of 5 for cost of computation, reflecting the necessity of a complex examination of all of the components of the matrix of majorities.

The opportunities for strategy under the ranked pairs rule are the same as for the Condorcet, simplified Dodgson, Nanson and Maximin rules (as long as truthful votes do not generate cycles), so its score for resistance to strategy is 8.9.

The ranked pairs rule is unique among the rules examined so far in possessing both Condorcet consistency and independence of clones. However, it is quite complex. It might also be faulted for its "lexicographic" treatment of majorities. Any number of majorities of 17 will be overruled to preserve one majority of 18.

The Weighted Condorcet Rule

Looking for a rule that would continue to possess both Condorcet consistency and independence of clones while not treating majorities lexicographically, I asked Thomas Zavist if he could suggest a way of weighting the rows and columns of a matrix of paired comparisons so that a rule like the Condorcet rule would be independent of clones. He suggested that this could be accomplished by weighting each row and column by the number of voters who ranked an option first. Thus the preliminary weighted Condorcet rule is the rule that each *order of the options* is scored by

$$\sum_{i=1}^{k-1} \sum_{j=i+1}^{k} w_i w_j C_{ij} \tag{13.2}$$

where w_i is the number of voters who rank in first place the option that is in position i for the order of the options that is being scored, and C_{ij} is the number of voters who rank the option in position i ahead of the option in position j. The limits on the double summation represent the top triangle of a square matrix. The option that is first in the order with the highest score is the winner.

The difficulty with the preliminary weighted Condorcet rule is that if an option receives no first-place votes, then its row and column of the matrix of paired comparisons receive no weight. As a consequence, if an option with no first-place votes is moved to any position in an order of the options, while leaving the order of the other options unchanged, there is no effect on the preliminary weighted Condorcet score.

This difficult can be resolved by sorting the options into "classes" in the following way: The first class consists of all options that have at least one first-place vote. The second class consists of all options that have at least one first-place vote after the first class has been deleted. The third class consists of all options that have at least one first-place vote after the first and second classes have been deleted, and so on until all options are in a class. Next, apply the preliminary weighted Condorcet rule to each class separately. Finally, integrate all the classes according to the following rule: First, place all options in the first class in the order in which the preliminary weighted Condorcet rule places them. Next, take the options in the second class and integrate them into the first class, beginning with the highest-ranked option in the second class, according to the following rule: If the first option in the second class beats all options in the first class, then it is placed ahead of all options in the first class. (Ties are resolved by a TBRO.) Otherwise it is placed immediately below the lowest-ranked option that beats it. Each of the remaining options is place in the highest position that is below the options in the second class that preceded it and below all options in the first class that beat it in paired comparisons. If there is a third class, it is handled as the second class was handled, now treating the first and second classes together as the first class was treated in integrating the second class, and so on until all classes are integrated. This is the weighted Condorcet rule.

The weighted Condorcet rule possesses mutual majority consistency (and therefore majority consistency), Smith consistency (and therefore Condorcet consistency), invariant loss consistency, resolvability, homogeneity, and independence of clones, but not Schwartz consistency, positive responsiveness, perfect-tie responsiveness, monotonicity, or combinativity.

The weighted Condorcet rule possesses mutual majority consistency (and therefore majority consistency) by the following argument. Let S be a set of options that are all ranked ahead of all other options by a majority of the voters, and let r be an order of the options that places r_1, an option not is S, first. If r_1 has any first-place votes, then it cannot be ranked first (within the first class of options) by the preliminary weighted Condorcet rule, by the argument that showed that the Condorcet rule possesses mutual majority consistency. And if r_1 has no first-place votes, then, since it loses to all options in S, it cannot be placed first at any subsequent stage of the weighted Condorcet rule.

The weighted Condorcet rule possesses Smith consistency (and therefore Condorcet consistency) for the following reason. If there is a set, S, of options that beat all other options in paired comparisons by majority rule, and r is a ranking

that places an option not in S first, then let i be the highest-ranked option in r and j be the highest-ranked option (by the preliminary weighted Condorcet rule) in the highest class that has options in S. If i and j are in the same class, then j must be ranked ahead of i by the weighted Condorcet rule because the Condorcet rule possesses Smith consistency. If the class of i is lower (farther from 1) than the class of j, then i is placed below j when the class of i is handled by the weighted Condorcet rule. If the class of i is higher than the class of j, then j is placed ahead of i when the class of j is handled, because, as the highest-ranked option in the first class with options in S, j must be placed above all other options. Thus r cannot be the order of the options that is chosen by the weighted Condorcet rule, and the weighted Condorcet rule possesses Smith consistency, and therefore Condorcet consistency.

A proof with essentially the same logic establishes that the weighted Condorcet rule possesses invariant loss consistency.

The weighted Condorcet rule possesses resolvability for the following reason. Within any class, a tie is resolvable by one vote by the argument that showed that the Condorcet rule possesses resolvability. If there is a tie between options in different classes, one vote resolves it in the process of integrating the classes. Thus the weighted Condorcet rule possesses resolvability.

The weighted Condorcet rule possesses homogeneity because, within each class, repeating a sequence of rankings k times multiplies all Condorcet scores by a factor of k, so that the weighted Condorcet rule yields the same order of options for each class. And since all paired comparisons are preserved, the integration of the classes yields the same final ordering.

To show that the weighted Condorcet rule possesses independence of clones one must first show that it will group clones together in the winning order of the options. Suppose on the contrary that in the order of options that wins by the weighted Condorcet rule, Q, x is immediately below options $y_1, y_2, \cdots, y_n$, which are not clones of x, and x', a clone of x, is immediately above options $y_1, y_2, \cdots, y_n$. Define Q' to be an order that is the same as Q except that x' has been placed between x and y_1. Define Q'' to be an order that is the same as Q except that x has been placed between y_n and x'. Since Q wins, it beats Q'. But the difference in the weighted Condorcet scores of Q and Q' is the same as the difference in the weighted Condorcet scores of Q and Q'', namely the sum of the majorities of x over $y_1, y_2, \cdots, y_n$. If this sum is positive, then Q'' beats Q, establishing a contradiction. If the sum is zero, then, since x and x' are clones, the TBRO that put x' ahead of $y_1, y_2, \cdots, y_n$ would also put x ahead of $y_1, y_2, \cdots, y_n$, establishing a contradiction again. Thus all clones must be grouped together in the winning order of the options under the weighted Condorcet rule.

Within the highest class that contains clones (call it class c), if a clone is introduced, the only voters who might rank that option first within class c are the voters who previously ranked one of the other clones first within class c. Thus the sum of the weights for the clones in class c is unaffected by the introduction of the

clone, and the weight for each non-clone is unaffected as well. The paired comparison between the new clone and any non-clone, x, is the same as the paired comparison between every other clone and x. And the new clone must be ranked in the midst of the other clones. The only effect of the introduction of a clone on preliminary weighted Condorcet scores of orders of the options comes from the comparisons of clones with other clones. For a given order of the clones among themselves, there is no differential impact on the relative scores of orders of the non-clones, with the set of clones inserted in any specified place. Thus within class c, if the winner was a clone before the introduction of a clone, the winner remains some clone. And if the winner was a non-clone, that option remains the winner. Thus the weighted Condorcet rule is independent of clones.

Example 13.5 shows that the weighted Condorcet rule lacks Schwartz consistency. For the weighted Condorcet rule, as for the regular Condorcet rule, w and x are tied winners, but the GOCHA set consists only of w.

Example 13.24 shows that the weighted Condorcet rule lacks monotonicity.

Example 13.24

1	1	1	1	1	1
w	x	y	z	x	z
x	y	z	w	w	y
y	z	w	x	y	w
z	w	x	y	z	x

For this election, the order $wxyz$ has a matrix of majorities of

	w	x	y	z
w	0	2	0	−2
x	−2	0	2	0
y	0	−2	0	2
z	2	0	−2	0.

The weights of w, x, y and z are 1/6, 1/3, 1/6 and 1/3 respectively, making the weighted sum of negative components of the upper triangle of the matrix of majorities − 1/9. But each of the orders $xyzw$, $yzwx$ and $zwxy$ yields the same weighted sum, so the result is a four-way tie among these orders. If the fifth voter interchanges the positions of options w and x, then the matrix of majorities for order $wxyz$ becomes

	w	x	y	z
w	0	4	0	−2
x	−4	0	2	0
y	0	−2	0	2
z	2	0	−2	0.

The weights of *w*, *x*, *y* and *z* become 1/3, 1/6, 1/6 and 1/3 respectively, so the weighted sum of negative components of the upper triangle becomes – 2/9. Order *yzwx* now has a matrix of majorities of

	y	*z*	*w*	*x*
y	0	2	0	– 2
z	– 2	0	2	0
w	0	– 2	0	4
x	2	0	– 4	0,

so its weighted sum becomes – 1/18, and it becomes the winner. Since the rise in the position of *w* caused it to change from a tied winner to a loser, the weighted Condorcet rule lacks monotonicity. It therefore also lacks positive responsiveness.

Example 13.7 shows that the weighted Condorcet rule lacks perfect-tie responsiveness. Under the weighted Condorcet rule, the winning order for Example 13.7 is *zxy*, for which the matrix of majorities is

	z	*x*	*y*
z	0	2	– 2
x	– 2	0	4
y	2	– 4	0,

with weights for options *x*, *y* and *z* of 0.4, 0.3, and 0.3 respectively. Thus the weighted Condorcet rule lacks perfect-tie responsiveness.

Example 13.8 shows that the weighted Condorcet rule lacks combinativity. By the weighted Condorcet rule, *x* wins for sequence **R** and for sequence **R′**, but *z* wins for sequence **R** + **R′**.

Turning to qualitative attractiveness, I give the weighted Condorcet rule a score of 3 for lucidity, equal to that of the estimated centrality rule and the lowest given to any rule. It score for ease of use is 5, like all other rules that ask for rankings. I give it a score of 2 for cost of computation, better than estimated centrality and worse than Condorcet, Dodgson and Young.

The opportunities for strategy under the weighted Condorcet rule are likely to be the same as for the Condorcet, simplified Dodgson, Nanson and Maximin rules when truthful votes do not generate cycles, so I give it a score of 8.9 for resistance to strategy.

The lack of monotonicity is a significant problem with the weighted Condorcet rule. It would be interesting to know whether there is some variation on the weighted Condorcet rule that possesses monotonicity.

The Schulze Method

Markus Schulze proposed a voting rule that is sometimes known as beatpath, sometimes as cloneproof Schwartz sequential dropping, and sometimes as variations on these names.[42] Since there are two quite different ways to determine the winner, Schulze suggests that it be called the Schulze method, so that the name will not appear to exclude either of these ways.[43]

The "beatpath" way to determine the winner of the Schulze method employs a transitive relation among options, defined as follows. A "path" from option x to option y is a sequence of options that begins with x and ends with y. A path is a "chain" if, for every pair, (i, j), of successive options in the path, $M(i, j) > 0$. Just as a chain is only as strong as its weakest link, the "strength" of a chain is the minimum of the majorities of the pairs of successive options in the chain.

If majority rule yields a transitive ordering of the options, then for every pair of options (x, y) for which $M(x, y) > 0$, there is a chain from x to y and there is no chain from y to x. But if x and y are in a cycle, then there is both a chain from x to y and a chain from y to x. If the cycle involves multiple options, then there can be many chains from x to y and many from y to x.

For any ordered pair of options (x, y), define $S(x, y)$ to be the strength of the strongest chain from x to y if there are any chains from x to y, and 0 if there are no chains from x to y. Define a relation between pairs of options, D (disqualifies), by $x\ D\ y$ if and only if $S(x, y) > S(y, x)$. The proof that follows shows that D is transitive.[44]

Suppose that $x\ D\ y$ and $y\ D\ z$. Then:

$$S(x, y) > S(y, x); \tag{13.3}$$

$$S(y, z) > S(z, y). \tag{13.4}$$

Since one path from x to z is the strongest chain from x to y followed by the strongest chain from y to z,

$$S(x, z) \geq \text{Min}[S(x, y), S(y, z)]. \tag{13.5}$$

Applying this idea to other pairs of links,

$$S(y, x) \geq \text{Min}[S(y, z), S(z, x)]; \tag{13.6}$$

$$S(z, y) \geq \text{Min}[S(z, x), S(x, y)]; \tag{13.7}$$

Consider separately the cases of $S(x, y) \geq S(y, z)$ and $S(x, y) < S(y, z)$. If

$$S(x, y) \geq S(y, z), \tag{13.8}$$

then in view of (13.4),

$$S(x, y) > S(z, y). \tag{13.9}$$

This says that one of the two items on the right side of (13.7) is greater than the left side of (13.7), so it must be the other item on the right side of (13.7) that is the smaller, so that

$$S(z, y) \geq S(z, x). \tag{13.10}$$

From (13.8), the second of the two items on the right side of (13.5) is the smaller, so that

$$S(x, z) \geq S(y, z). \tag{13.11}$$

Combining (13.11), (13.4) and (13.10),

$$S(x, z) \geq S(y, z) > S(z, y) \geq S(z, x), \tag{13.12}$$

which means that $x \, D \, z$.

Now consider the case of $S(x, y) < S(y, z)$. Reversing the two sides,

$$S(y, z) > S(x, y), \tag{13.13}$$

which, in view of (13.3), implies that

$$S(y, z) > S(y, x). \tag{13.14}$$

This says that one of the two items on the right side of (13.6) is greater than the left side of (13.6), so it must be the other item on the right side of (13.6) that is the smaller, so that

$$S(y, x) \geq S(z, x). \tag{13.15}$$

From (13.13), the second of the two items on the right side of (13.5) is the greater, so that

$$S(x, z) \geq S(x, y). \tag{13.16}$$

Combining (13.16), (13.3), and (13.15),

$$S(x, z) \geq S(x, y) > S(y, x) \geq S(z, x), \tag{13.17}$$

which again means that $x \, D \, z$, so that D is transitive.

The "beatpath" definition of the Schulze method is that the winning options in an election are the options that are not disqualified by any other option by the relation *D*. Since *D* is transitive, there is at least one winning option. If there is more than one winning option, the tie among the winning options is broken by a TBRO, constructed in the way described in the discussion of the ranked pairs rule.

One way to determine the beatpath winner is to calculate the strength of every path from every option to every other option and determine which options are not disqualified. Schulze offers an efficient algorithm for doing this.[45] But it is not necessary to use such an algorithm to determine the winner under the Schulze method.

When paired comparisons by majority rule establish a transitive order of the options, the Schulze method selects the dominant option, since every option in such a transitive order disqualifies every option below it. (No chain can go up a transitive order.) When the Smith set for an election contains more than one option, the Schulze method selects something from the Smith set, since every option inside the Smith set disqualifies every option outside it. (No chain can enter the Smith set from outside.)

In fact, if the GOCHA set is smaller than the Smith set, then the winner by the Schulze method must come from the GOCHA set. The reason is as follows. For some election, define *G* as the GOCHA set, *S* as the set of options that are in the Smith set but not in *G*, and *E* as the set of all other options (which may be the empty set). By the definition of the Smith set, all options in *G* and *S* beat all options in *E*. If all the options in *G* beat all the options in *S*, then the options in *S* would be in *E*. If any of the options in *S* beat any of the options in *G*, then at least some of the options in *S* would be in *G*. If none of the options in *G* beat any of the options in *S*, then the options in *S* would be in *G*. This means that if *S* is not empty, then in paired comparisons between an element of *G* and an element of *S*, there must be some cases in which the element of *G* wins, some in which there is a tie between the element of *G* and the element of *S*, and none in which the element of *S* wins. Therefore there is at least one chain that goes from an element of *G* to and element of *S* and no chain that goes from any element of *S* to any element of *G*. Therefore any *s* in *S* to which a chain from an element of *G* goes is disqualified, as are all elements of *S* that are beaten by *s*, and all that are beaten by these, and so on until no more are added. If there were elements of *S* that could not be disqualified in this way, then they would be an undominated set, and any minimal undominated subset of them would be in *G* and not in *S*. Thus all elements of *S* are disqualified, implying that the winner under the Schulze method must be an element of the GOCHA set. This means that it is generally easiest to determine the winner under the Schulze method by first determining the GOCHA set and then determining which elements of the GOCHA set are disqualified.

One can finish the task of identifying the winner under the Schulze method by determining the strengths of all of the chains within the GOCHA set and thereby determine which options are disqualified, but there is a method of successive eliminations that is simpler. Let m_1 be the minimum of the majorities between pairs of options in the GOCHA set. If some option, *x*, has a chain to some other

option, y, of strength m_1, then x will be disqualified by y unless it happens that the chain from y to x also has strength m_1. The disqualifications will be the same as if every majority of m_1 is replaced by a majority of 0. As long as some majorities in the GOCHA set are greater than m_1, one can replace all majorities of m_1 with majorities of 0 and identify a new GOCHA set in which the winning option must lie. Any options that were in the previous GOCHA set and not in the new GOCHA set are disqualified.

Let m_2 be the minimum of the majorities between pairs of options in the new GOCHA set. As long as some majorities in the new GOCHA set are greater than m_2, one can replace all majorities of m_2 with majorities of 0 and identify a new GOCHA set in which the winning option must lie. This process continues until all majorities in the GOCHA set are equal. All options then in the GOCHA set are the winners. These are the options that do not disqualify each other and do disqualify all options not in the final GOCHA set. Thus they are the same as the winners under the beatpath description of the Schulze method.

The Schulze method possesses Schwartz consistency, as shown above (and therefore Smith consistency, Condorcet consistency and majority consistency), mutual majority consistency, invariant loss consistency, monotonicity, homogeneity, and independence of clones. It lacks positive responsiveness, perfect-tie responsiveness, and combinativity. I do not know whether the Schulze method is resolvable.[46]

The Schulze method possesses mutual majority consistency because if there is a set of options, S, such that a majority of voters rank all options in S ahead of all other options, then no option outside of S could be in the Smith set. Since the Schulze method possesses Smith consistency, it selects a winner from S, so it possesses mutual majority consistency. Similarly, The Schulze method possesses invariant loss consistency because an option that loses to all other options cannot be in the Smith set.

The Schulze method possesses monotonicity by the following argument. Suppose that the sequence of rankings $\mathbf{R}'$ is an incremental upward movement for x relative to the sequence $\mathbf{R}$, and that x is among the winning options for $\mathbf{R}$. For any chain from x to an option that is disqualified by x, the incremental upward movement either increases the strength of the chain or has no effect on it, while for any chain to x, the incremental upward movement either reduces the strength of the chain or leaves it unaffected. Thus under rankings $\mathbf{R}'$, x continues to disqualify all options that it disqualified under $\mathbf{R}$. For any options that were not disqualified by x under $\mathbf{R}$ (which would be options that were tied with x for winning status), the incremental upward movement might cause x to now disqualify them and could not cause them to disqualify x. Thus x continues to be a winning option.

The Schulze method possesses homogeneity because replicating a sequence of rankings n times causes all $M(i, j)$ to be multiplied by n and has no effect on the pattern of disqualifications.

The Schulze method is independent of clones by the following argument. If c', a clone of c, is added, then for every x that is not a clone of c, $M(c', x) = M(c, x)$, so that c' disqualifies x if and only if c disqualifies x, and x disqualifies c' if and only

if x disqualifies c. Therefore if a clone is a winner before c' is added, a clone remains among the winner after c' is added, and any x that is a winner before c' is added remains a winner after c' is added.

Example 13.6 shows that the Schulze method lacks positive responsiveness. By the Schulze method, the election is a tie between x and y both before and after the incremental upward movement for x.

Example 13.7 shows that the Schulze method lacks perfect-tie responsiveness. The addition to the perfect tie of a vote ranking x first leaves a tie between x and z by the Schulze method.

Example 13.8 shows that the Schulze method lacks combinativity. It awards the elections for sequences $\mathbf{R}$ and $\mathbf{R}'$ to x and the election for sequence $\mathbf{R} + \mathbf{R}'$ to z.

Turning to measures of qualitative attractiveness, I give the Schulze method a score of 4 for lucidity, a notch below the Young rule and the ranked pairs rule. I give it a score of 5 for ease of use, like all rules that require rankings. I give it a score of 4 for cost of computation, a notch below the ranked pairs rule, because of the relative difficulty of identifying the GOCHA set.

The opportunities for strategy under the Schulze method are likely to be the same as for the Condorcet, simplified Dodgson, Nanson and Maximin rules when truthful votes do not generate cycles, so its score for resistance to strategy is 8.9.

The Schulze method is substantially similar to the ranked pars rule, but definitely different since it possesses Schwartz consistency and the ranked pairs rule does not. A choice between the two could reasonably be based on whether one believes that a vote processing rule ought to possess Schwartz consistency.

The Alternative Smith and Alternative Schwartz rules

The ranked pairs rule and the Schulze method are noticeably more susceptible to strategy than the alternative vote. What gives the alternative vote its greater resistance to strategy is the fact that it eliminates options based on first-place votes rather than paired comparisons. Strategants cannot reduce the number of first-place votes of a leading option that they are seeking to defeat. This suggests that, to create rules that satisfy as many non-qualitative criteria as the ranked pairs rule and the Schulze method, while also having as much resistance to strategy as the alternative vote, one might begin by eliminating all options that are not in the Smith set or the GOCHA set, and then eliminate the option with the fewest first-place votes. One would then form the Smith set or GOCHA set with respect to comparisons among the remaining options and, if this set contained more than one option, again delete the option with the fewest first place votes. Ties would be resolved by a TBRO. I call these rules alternative Smith and alternative Schwartz, because they combine the elimination rule from the alternative vote with the concepts of Smith consistency and Schwartz consistency.

The alternative Schwartz rule possesses Schwartz consistency (and therefore Smith consistency, Condorcet consistency and majority consistency), mutual majority consistency, invariant loss consistency, homogeneity, and independence

of clones. It lacks monotonicity (and therefore positive responsiveness), resolvability, perfect-tie responsiveness, and combinativity.

The alternative Schwartz rule possesses Schwartz consistency (and therefore Smith consistency, Condorcet consistency and majority consistency) by the way it is defined. It possesses mutual majority consistency because if the options in set S have a mutual majority over all other options, then no option that is not in S could be in the GOCHA set. It possesses invariant loss consistency because no option that lost to all others in paired comparisons could be in the GOCHA set. It possesses homogeneity because replicating a sequence of rankings k times preserves all equalities and inequalities with respect to comparisons of the numbers of first-place votes and the sizes of majorities in paired comparisons. It possesses independence of clones because the introduction of a clone has no effect on whether any non-clone is in the GOCHA set, no effect on whether the GOCHA set contain some clone, and no effect on whether the last clone will be eliminated before or after the last non-clone.

Example 13.10 shows that the alternative Schwartz rule lacks monotonicity and therefore lacks positive responsiveness.

Example 13.25 shows that the alternative Schwartz rule lacks resolvability.

Example 13.25

2	1	2	1
w	*x*	*y*	*z*
x	*z*	*z*	*x*
y	*w*	*w*	*w*
z	*y*	*x*	*y*

For this election, the matrix of majorities is

	w	*x*	*y*	*z*
w	0	2	2	−2
x	−2	0	2	0
y	−2	−2	0	2
z	2	0	−2	0.

The GOCHA set is all four options, and there is a tie between x and z for fewest first-place votes. If z is eliminated, then w wins since it beats both x and y in paired comparisons. But if x is eliminated, then there is a symmetric three-way tie among w, y, and z. An extra vote for z would guarantee that x rather than z would be eliminated in the first round, but then there would be a tie between w and y in the second round for fewest first-place votes, and if w is eliminated, then y rather than z wins. Thus there is no way to produce a win for z by adding a vote, so the alternative Schwartz rule lacks resolvability.

Example 13.7 shows that the alternative Schwartz rule lacks perfect-tie responsiveness. Example 13.8 shows that it lacks combinativity.

By similar reasoning, the alternative Smith rule possesses the same properties as the alternative Schwartz rule, except that Example 13.5 shows that it lacks Schwartz consistency.

In terms of qualitative measures, I give the alternative Schwartz rule a score of 4 and alternative Smith rule scores of 5 for lucidity, reflecting a judgment that the GOCHA set is a more complex concept than the Smith set. I give both rules scores of 5 for ease of use (like all other rules that employ rankings). For cost of computation, I give the alternative Smith rule a score of 5 and the alternative Schwartz rule a score of 4, reflecting a judgment that the GOCHA set is noticeably more difficult to identify than the Smith set.

It is very difficult to strategize successfully against the alternative Schwartz and alternative Smith rules. In the typical case where there is no cycle, the strategants must both create a cycle and arrange for the otherwise winning option to have the fewest first-place votes. If non-strategic voting produces a transitive ranking of three options *xyz* and *x* has more than a third of the first-place votes, then no strategy can succeed. If there is a transitive ranking of *xyz* and *x* has fewer than a third of the votes, then it may be possible to devise a strategy in which some voters whose true rankings are *yxz* change to *yzx* to create a cycle, while other voters change from *yzx* to *zyx* or vice versa, so that *x* will have the fewest first-place votes, but devising these strategies is more challenging than devising strategies for other voting rules.

Among the 87 elections in my sample, there were just two for which I found strategies that changed the outcome when the alternative Schwartz or alternative Smith rule was used. Election 13 involved a semi-cycle among four options at the top of the ordering, as shown below.

	w	*x*	*y*	*z*
w	0	4	2	0
x	−4	0	1	0
y	−2	−1	0	1
z	0	0	−1	0

By the alternative Schwartz rule, this election is won by *w*, since it constitutes the only minimal undominated set and is therefore the only element of the GOCHA set. The Smith set, on the other hand, consists of all four options, so the one with the fewest first place votes must be eliminated. This turns out to be *x*. The remaining three options form the new Smith set, of which *w* has the fewest first-place votes after those of *x* are distributed. After *x* is eliminated, *y* beats *z* by one vote.

The simplest strategy against the alternative Schwartz rule is for one voter who finds *z* more attractive than *y* to rank *z* ahead of *w*, so that *w* will no longer be undominated and all four options will be in the GOCHA set. Then *y* wins. The

simplest strategy against the alternative Smith rule is for one voter who prefers *w* to *y* to rank *w* ahead of *z*. Then *w* is dominant and wins.

Election 26 is the only other election in my sample that was subject to strategy by the alternative Schwartz or alternative Smith rule. It also involved a semi-cycle at the top, this time among three options, as shown below.

	x	*y*	*z*
x	0	5	0
y	−5	0	3
z	0	−3	0

The only element in the GOCHA set is *x*, so it wins by the alternative Schwartz rule. The Smith set consists of all three options, so the one with the fewest first-place votes is eliminated. This turns out to be *y*, making the election a tie between *x* and *z* when the alternative Smith rule is used.

When the alternative Schwartz rule is used, there is a two-part strategy available to those who prefer *z* to *x*. If some of them place *y* ahead of *x*, so that *x* no longer beats *y*, while others place *y* below *z*, so that *z* will beat *y* and the first action will not have made *y* dominant, then the election will be a tie, and they will have a 50-50 chance of electing the option they want. I have not found a strategy that succeeds for this election when the alternative Smith rule is used.

In calculating the overall resistance to strategy, I adjust the computed figure half way to 10, for the calculation of the resistance to strategy of the alternative Schwartz rule, to allow for the fact that the strategy achieved only a tie. With this adjustment, the score of the alternative Schwartz rule for resistance to strategy is 9.846, and the score for the alternative Smith rule is 9.898. Inquiries with other data sets would be needed to be confident that there is any reliable difference in the resistance to strategy of the two rules.

Thus using first-place votes as an elimination criterion after reducing the contenders to the GOCHA set or the Smith set does produce rules that retain the consistency properties while greatly increasing resistance to strategy, compared to other rules that have the consistency properties. In the process, monotonicity and resolvability are lost.

Cardinal-Weighted Pairwise Comparisons

James Green-Armytage has suggested the following voting method, to limit the influence of strategy.[47] Have each voter report a score for each option on a scale such as 0 to 100. To count the votes, begin by using the rankings implied by the scores to reduce the set of contending options to the Smith set. If the Smith set contains only one option it is the winner. If the Smith set contains more than one option, then the final ordering of the options must be inconsistent with one or more

paired comparisons. To decide which one or more paired comparisons are to be disregarded, the paired comparisons are ordered, but not in the usual way. Re-scale each voter's vote so that the lowest option in the Smith set has a score of 0 and the highest a score of 100.[48] (The re-scaling assumes that there is some variation in the scores given to the options in the Smith set. If not, the vote does not contribute to choosing within the Smith set.) Then each paired comparison is given a score equal to the sum, over voters on the winning side of that comparison, of the difference between the re-scaled scores of the options in the comparison. Finally, some method such as ranked pairs or the Schulze method is used to decide which of the scored paired comparisons will be followed and which will be disregarded in the final ranking. Green-Armytage calls his system cardinal-weighted pairwise comparisons, or cardinal pairwise for short.

The central idea of cardinal-pairwise is that if the choice of a winning option must be inconsistent with one or more paired comparison, the claim of a particular comparison to be followed in the final ordering should be based on how important that comparison is to those who support it, as measured by the differences in the scores that they give to the options in the comparison.

The cardinal pairwise system satisfies all of the consistency properties other than Schwartz consistency, for the same reasons that the alternative Smith rule satisfies them. If Schwartz consistency is desired, it can be achieved by specifying that the options will be initially reduced to the GOCHA set. It is homogeneous because replicating all votes changes nothing on which the outcome depends. It is independent of clones when every voter gives uniform scores to any set of clones, for the same reason that the ranked pairs rule is independent of clones. It is probably resolvable because of the continuity of scores. Green-Armytage provides an example that shows that the cardinal pairwise system is not monotonic.[49]

In terms of qualitative attractiveness, I give the cardinal pairwise system a score of 3 for lucidity, 5 for ease of use, and 3 for cost of computation. To estimate its resistance to strategy, one would need a sample of elections in which voters scored the options.

Green-Armytage argues that by ensuring that the final outcome will be consistent with the paired comparisons about which voters feel most strongly, and by requiring strategants to give less weight to one paired comparison if they wish to give more weight to another, the cardinal pairwise system will make successful strategies more difficult. I find this claim intriguing and worthy of further inquiry, but I have been unable to decide whether to endorse it.

Summary

Table 13.1 displays the properties of the rules that have been discussed. Table 13.2 divides 22 of the 23 rules mentioned in Table 13.1 into five categories of supportability. (Cardinal pairwise is not evaluated because it has not been studied enough to compare it with the others.)

Table 13.1 Properties of vote-processing rules

	Consistency							Respon-siveness			Stability				Qualitative attractiveness			
	Universal domain	Majority consistency	Mutual maj. consistency	Condorcet consistency	Smith consistency	Schwartz consistency	Invar. loss consistency	Positive responsiveness	Resolvability	Perfect-tie responsiveness	Monotonicity	Homogeneity	Combinativity	Independence of clones	Lucidity	Ease of use	Cost of computation	Resistance to strategy
Plurality	o	x	[a]	x	x	x	x	x	x	x	x	x	x	x	10	10	10	6.3
Plurality with UD	x	x	o	o	o	o	o	o	x	x	x	x	x	o	6	5	9	6.3
Approval voting	o	x	x	x	x	x	x	x	x	x	x	x	x	x	9	7	9	3.9
Approval with UD	x	o	o	o	o	o	o	o	x	x	x	x	x	x[b]	6	4	8	3.9
Range voting	x	o	o	o	o	o	o	x	x	x	x	x	x	x[b]	10	5	8	4.0
Borda	x	o	o	o	o	o	x	x	x	x	x	x	x	o	9	5	7	4.6
Condorcet	x	x	x	x	x	o	x	o	x	o	x	x	o	o	6	5	3	8.9
Two-ballot majority	x	x	o	o	o	o	x	o	o	o	o	x	o	o	9	8	8	8.1
Alternative vote	x	x	x	o	o	o	x	o	o	o	o	x	o	x	8	5	6	9.7
Dodgson	x	x	o	x	o	o	o	o	o[c]	o	o	o	o	o	6	5	3	?
Simplified Dodgson	x	x	o	x	o	o	o	o	x	o	x	x	o	o	8	5	7	8.9
Nanson	x	x	x	x	x	o	x	o	?	o	o	x	o	o	7	5	6	8.9
Bucklin	x	x	x	o	o	o	o	o	x	o	x	x	o	o	8	5	6	5.8
Copeland	x	x	x	x	x	o	x	o	o	o	x	x	o	o	8	5	7	9.4
Black	x	x	o	x	o	o	x	x	x	x	x	x	o	o	7	5	7	7.8
Maximin	x	x	o	x	o	o	o	o	x	o	x	x	o	o	7	5	7	8.9
Coombs	x	o	o	o	o	o	x	o	o	o	o	x	o	o	7	5	6	?
Estimated centrality	x	?	o	o	o	o	?	?	x[d]	?	?	x	o	o[d]	3	5	1	?
Young	x	x	o	x	o	o	o	o	x	o	x	x	o	o	5	5	3	?
Ranked pairs	x	x	x	x	x	o	x	o	x	o	x	x	o	x	5	5	5	8.9
Weighted Condorcet	x	x	x	x	x	o	x	o	x	o	o	x	o	x	3	5	2	8.9
Schulze	x	x	x	x	x	x	x	o	?	o	x	x	o	x	4	5	4	8.9
Alternative Schwartz	x	x	x	x	x	x	x	o	o	o	o	x	o	x	4	5	4	9.8
Alternative Smith	x	x	x	x	x	o	x	o	o	o	o	x	o	x	5	5	5	9.9
Cardinal pairwise	x	x	x	x	x	x	x	?	x[d]	?	o	x	?	x[b]	3	5	3	?

x means that a rule possesses a property; o means that it does not. Higher scores denote greater qualitative attractiveness.

a. Not defined.

b. Possesses the property when clones are redefined to reflect the domain of the rule.

c. Resolvable by two votes when not resolvable by one.

d. Conjecture.

Table 13.2 A classification of vote-processing rules by supportability

Not supportable	Arguably inferior to maximin	Supportable if ranking is infeasible	Supportable if a matrix of majorities is uncalculable	Supportable if a matrix of majorities is calculable
Borda	Condorcet	Plurality	Alternative vote	Maximin
Range	Simp. Dodgson	Approval		Ranked pairs
Dodgson	Nanson	Two-ballot maj.		Schulze method
Copeland	Bucklin			Alt. Schwartz
Coombs	Black			Alt. Smith
Est. centrality	Young			
	Wt. Condorcet			

Six of the rules are classified as unsupportable. All of these have defects that are so serious as to disqualify them from consideration. The Borda rule lacks Condorcet consistency and is highly susceptible to strategizing. Because of this, as Borda himself realized, the Borda rule can be recommended only for situations where voters are disinterested judges, where no one believes that the voters might vote strategically. It cannot be recommended for situations where voters are self-interested advocates. Range voting has the same limitation and should not be used where voters are self-interested advocates. When voters can be relied upon to act as disinterested judges, range voting has the virtue that the Borda rule does not, of not presuming that voters perceive the options to be equally spaced in merit.

The Dodgson rule was not actually advocated by Dodgson and has received little if any support from others. If one found this rule attractive, one would presumably find the simplified Dodgson rule more attractive, for its homogeneity and monotonicity.

The most important objection to the Copeland rule is its lack of resolvability. It does not provide an answer to the question of which option wins for any simple asymmetric cycle among three options. This is reason enough to reject it.

The Coombs rule has a rationale only for voters' preferences that are single-peaked in one dimension, a condition that cannot be presumed to hold in general. The practice of eliminating the option with the most last-place votes makes the Coombs rule highly susceptible to cloning. It also lacks Condorcet consistency. For these reasons it does not deserve serious consideration.

The estimated centrality rule is so complex that the only computer program that has been written to compute it dealt only with the case of three options. It is also highly susceptible to strategizing. For this reason, even if the difficulties of computing could be resolved, it would be suitable only for situations where the voters are judges rather than self-interested advocates.

The seven rules in the next category are all sophisticated ways of processing rankings. Each has a coherent rationale. However, each, for its own reason, is arguably less attractive than the maximin rule.

The Condorcet rule could have an unmanageable cost of computation, particularly when the number of options is great. However, the evidence from Chapter 9 suggests that even when the number of options is great, cycles are quite rare. Thus is it not clear that the potential for a high cost of computation should exclude the Condorcet rule. When compared to the maximin rule, the Condorcet rule has the additional property of Smith consistency. However, the example that showed that the maximin rule lacked this property involved a rather contrived example of a cycle within a cycle. Neither rule possesses independence of clones, but the Condorcet rule can be affected by a single replication of one option, while the maximin rule is susceptible to clones only when there is a cycle among the clones. Thus the advantage of the Condorcet rule in Smith consistency is arguably insufficient to justify its potentially high cost of computation and its added susceptibility to cloning.

The simplified Dodgson rule is shown in Table 13.1 to have the same formal properties as the maximin rule. However, under the simplified Dodgson rule, the number of different options to which a given option loses counts against it, and this invites a strategy in the form of replicating the options that one favors. While the maximin rule is not independent of clones, its susceptibility to cloning is confined to cases in which clones are involved in cycles. Thus the simplified Dodgson rule is arguably inferior to the maximin rule.

The Nanson rule can be criticized for its lack of monotonicity, a fault that occurred in all examined rules that operate through successive eliminations. More fundamentally, rules that rely on successive eliminations use a minimum of information at the final stage of processing and thereby fail to be sensitive to some information.

The Bucklin system lacks Condorcet consistency and is quite a bit more susceptible to strategy than some competing rules.

The Black rule is unique among rules that possess Condorcet consistency, both in its possession of positive responsiveness and in its possession of perfect-tie responsiveness. However, by combining two different approaches, it sacrifices coherence. It also provides a general incentive for cloning whenever there is a cycle. For these reasons it is less attractive than the maximin rule.

The Young rule has the same set of investigated formal properties as the maximin rule. It yields the same result as the maximin rule whenever every potentially winning option has as many last-place votes as its worst defeat in paired comparisons. But the Young rule is potentially very costly to calculate. On this basis, it seems appropriate to eliminate it in favor of the maximin rule.

The weighted Condorcet rule scores very low in both lucidity and cost of computation, and it lacks monotonicity, which could reasonably concern potential users.

Three vote-processing rules are classified as supportable if it is not feasible to require voters to rank options. The fact that these rules are not ranking-based gives them significant advantages in terms of the effort required of voters.

The plurality rule requires the least effort on the part of voters and has the further advantage of familiarity. However, the possibility that supporters of an idea will split their votes among similar options provokes a great deal of organizing effort. The appropriateness of the selected option is often called into question when it receives less than a majority of the votes. For these reasons the plurality rule is attractive only when simplicity is paramount. It ought to be replaced whenever voters can tolerate something more complex.

The least complex alternative to plurality is approval voting. The voting process is the same, except that voters are allowed to give a vote to more than one option. Counting the votes is also simple and straightforward. Approval voting overcomes the most obvious difficulty of plurality, the necessity imposed on voters of selecting a single option to support when there are several similar ones that could win if their supporters united and will lose if they do not. Approval voting would be an ideal method of voting if it corresponded to the thinking of voters about options, that is, if voters categorized options unproblematically as either worthy of approval or not. However, to the extent that voters think about options in terms of their relative merit compared to other options, approval voting will require voters to make strategic calculations as to where in their rankings they wish to cut off their approval. Therefore the supportability of approval voting is contingent on voters being intolerant of more complex possibilities.

An alternative way of overcoming some of the limitations of plurality is two-ballot majority. The act of voting under two-ballot majority is no more difficult than under plurality, except that two acts of voting may be required. And if an option receives a majority in the first round of voting, the second round is not needed. The first round of voting serves to inform less involved voters as to which options deserve their serious consideration. The second round insures that the selected option receives the imprimatur of a majority of votes against *some* alternative option. However, two-ballot majority does not overcome all problems of vote splitting, and it lacks many of the investigated formal properties. Both approval and two-ballot majority represent significant improvements over plurality, and they are probably as attractive as rules can be if it is not feasible to have voters rank options.

If it is feasible to require voters to rank options, then much more sophisticated processing is possible. However, it is conceivable that it would be feasible to require voters to rank options but not feasible to require vote-processors to produce a matrix of majorities. In this event, the alternative vote is supportable. It has the advantages of being currently used for some purposes and of being reasonably easy to understand and to count. But it has serious weaknesses as well. The alternative vote lacks both Condorcet consistency and monotonicity. Its practice of eliminating the option with the fewest first-place votes has the unfortunate

potential of eliminating natural compromise options. While the alternative vote offers significant advantages over plurality, one might hope that if it is feasible to have voters supply rankings of the options, it would then be possible to process those rankings in a more sophisticated fashion than the alternative vote.

The final category is rules that are supportable if their complexity in the act of voting and the cost of calculating a matrix of majorities are both tolerable. There are five rules in this category, differing in the amount of complexity that they require voters and vote-processors to tolerate.

Among rules that require a matrix of majorities, the maximin rule is extremely simple and yet very powerful. The rule can be explained in a sentence of just 11 words: The winner is the option whose worst loss is least bad. Only a very brief examination of the matrix of majorities is needed to determine the winner. While the maximin rule is not entirely independent of clones, it is independent of clones that come only in pairs, and its susceptibility to cloning occurs only with cycles among the clones. It lacks some consistency properties, but the frequency with which such cases arise may be low enough to make these deficiencies tolerable, in light of the relative simplicity of the rule.

The ranked pairs rule has all of the consistency properties other than Schwartz consistency. And whether Schwartz consistency is desirable in a rule that has Smith consistency can be questioned. The issue can be illustrated with a matrix of paired comparisons that arose in the discussion of the alternative Schwartz and alternative Smith rules:

	x	*y*	*z*
x	0	5	0
y	−5	0	3
z	0	−3	0

If option *y* were not present, this election would be a tie between *x* and *z*. When *y* is introduced and beats *z* while losing to *x* in paired comparisons, what does this imply? The Schwartz-consistency interpretation is that *y* reveals that *z* can be beaten and *y* is also beaten, so *x* as the only unbeaten candidate must be the winner. The Smith-consistency interpretation is that with the introduction of *y* there are three options rather than two in the smallest set such that all options in the set beat all options outside it. Some cycle-breaking rule must be used to choose among *x*, *y*, and *z*. I find the Smith-consistency interpretation more attractive, so I do not mind the lack of Schwartz consistency in the ranked pairs rule.

The Schulze method (in its sequential dropping algorithm) shares with the ranked pair rule a focus on ordering paired comparisons. But the Schulze method drops the weakest at each stage where the ranked pairs rule embraces the strongest that can be accepted in light of earlier commitments. The result is the same in most cases, but the Schulze method possesses Schwartz consistency while the

ranked pairs rule does not. A choice between these two can reasonably be based on how one feels about Schwartz consistency.

The three rules discussed so far in this final category are approximately equal in resistance to strategy and not nearly as resistant to strategy as the alternative vote. To devise a rule that would have the consistency properties of ranked pairs and Schulze while achieving a resistance to strategy that rivaled that of the alternative vote, I devised the alternative Schwartz and alternative Smith rules. Based on the sample of elections considered here, these rules do appear to have a greater resistance to strategy than any others. The alternative Schwartz rule offers Schwartz consistency if one wants it; the alternative Smith rule is available for those who prefer not to have Schwartz consistency. These rules lack monotonicity, but it possible that people choosing a vote processing rule would regard the frequency with which that anomaly arises as sufficiently low to be acceptable in view of the benefits that these rules offer in increased resistance to strategy.

Whether the cardinal pairwise method can compete with or surpass these five rules must await further inquiry.

If voters and vote counters have only a slight tolerance for complexity, the maximin rule is the one they would reasonably choose. If they have more tolerance for complexity, then they would reasonably choose alternative Schwartz or alternative Smith if resistance to strategy is important and slight deviations from monotonicity are tolerable. Otherwise they would choose ranked pairs or Schulze, with the choice between the two rules in either pair dependent on the value attached to Schwartz consistency.

The next chapter considers the problem of vote-processing rules for selecting one option from a continuum of one or more dimensions.

Notes

1 Steven J. Brams and Peter C. Fishburn, *Approval Voting* (Boston, 1983), p. 1.
2 For a fuller discussion of Borda's work, see Duncan Black, *The Theory of Committees and Elections* (Cambridge, 1958), pp. 156–9.
3 Ibid., p. 178.
4 Ibid., p. 179.
5 Ibid., p. 157.
6 Ibid., pp. 157–8.
7 Ibid., pp. 61–4.
8 Ibid., p. 180.
9 Ibid., pp. 179–80.
10 Ibid., pp. 160–78.
11 H. Peyton Young, "Condorcet's Theory of Voting," *American Political Science Review* 82 (1988): 1231–44.
12 J. Kemeny, "Mathematics without Numbers," *Daedalus* 88 (1959): 571–91.
13 I am indebted to James Green-Armytage for this example.

14 For a fuller discussion of two-ballot majority, see Robert Elgie, "Two-Ballot Majority Electoral Systems," *Representation* 34 (1997): 89–94.
15 Black, *The Theory of Committees and Elections*, pp. 189–213.
16 Ibid., p. 227.
17 Ibid., p. 226.
18 Ibid., p. 227.
19 Ibid., pp. 227–30.
20 Peter C. Fishburn, *The Theory of Social Choice* (Princeton,1973), pp. 172–3.
21 Black, *The Theory of Committees and Elections*, p. 186.
22 Clarence Hoag and George Hallett, *Proportional Representation* (New York, 1926), pp. 491–2.
23 Ibid., p. 486.
24 Ibid., pp. 486–7.
25 Leo Goodman, "On Methods of Amalgamation," in R.M. Thrall, C.H. Coombs, and R.L. Davis (eds), *Decision Processes* (New York, 1954), p. 42.
26 Black, *The Theory of Committees and Elections*, pp. 58–9.
27 Ibid., p. 66.
28 Ibid., p. 175.
29 Ibid., p. 69.
30 Ibid., pp. 69–70.
31 Clyde Coombs, *A Theory of Data* (New York, 1964), especially Chapter 5, pp. 80–121.
32 Ibid., p. 193.
33 Ibid., pp. 85–6.
34 Ibid., pp. 397–9.
35 Ibid., p. 399.
36 I.J. Good and T.N. Tideman, "From Individual to Collective Ordering Through Multidimensional Attribute Space," *Proceedings of the Royal Society of London, A* 347 (1976): 371–85.
37 H.P. Young, "Extending Condorcet's Rule," *Journal of Economic Theory* 16 (1977): 335–53.
38 T.N. Tideman, "Independence of Clones as a Criterion for Voting Rules," *Social Choice and Welfare* 4 (1987): 185–206.
39 T.M. Zavist and T.N. Tideman, "Complete Independence of Clones in the Ranked Pairs Rule," *Social Choice and Welfare* 6 (1989): 167–73.
40 Ibid.
41 This alternative definition is developed in Ibid., p. 172.
42 M. Schulze, "A New Monotonic and Clone-Independent Single-Winner Election Method," *Voting Matters* 17 (2003): 9–19 (www.mcdougall.org.uk/VM/MAIN.HTM).
43 Ibid., p. 9
44 The proof is based on one by Schulze in Ibid., pp. 10–11.
45 Ibid., p. 11.
46 Schulze (Ibid., p. 13) offers a proof that the Schulze method satisfies a condition that he call resolvability, but resolvability as defined here is a different condition.
47 J. Green-Armytage, "Cardinal-Weighted Pairwise Comparison," *Voting Matters* 19 (2004): 6–13 (www.mcdougall.org.uk/VM/MAIN.HTM).
48 Green-Armytage regards this step as optional, but it seems essential to me.
49 See http://fc.antioch.edu/~james_green-armytage/vm/cwp_criteria.htm#_ftn4.

This page intentionally left blank

Chapter 14

Vote-Processing Rules for Selecting One Option from a Continuum of One or More Dimensions When Votes Have Predetermined Weights

Introduction

There are two things about continuums that make vote-processing rules for choices from continuums different from vote-processing rules for choices from sets of discreet options: The number of options that can easily be compared is very large (potentially infinite), and there is generally reason to expect a particular structure for the preferences of voters. This chapter explores the consequences of these differences for vote-processing rules, first for choices from one-dimensional continuums and then for choices from continuums of two or more dimensions.

Vote-Processing Rules for Choices from One-Dimensional Continuums

A choice from a one-dimensional continuum can be exemplified by a decision like, "What will the speed limit on this road be?" or "How much should our community spend this year on parks?" As was mentioned in Chapter 9, when the options lie in a one-dimensional continuum and voters' preferences are single-peaked, there is a majority-rule equilibrium at the ideal point of the median voter when the number of voters is odd, and there is a majority-rule equilibrium everywhere between the two median voters when the number of voters is even. This suggests the possibility of using the "unidimensional median choice" (UMC) rule when options lie in a one-dimensional continuum:

> *Definition*: Unidimensional median choice rule: When a decision is to be made from a one-dimensional continuum, each voter shall state the option for which he or she wishes to vote. From these statements the median or medians shall be determined. If the median is unique, it shall be declared the winning option; if there are two median options that differ, then a random process shall be used to choose between them.

If voters' preferences are single-peaked, then every voter will have an incentive to vote for his or her first choice. A voter who is the median voter (when reporting his or her preference honestly) gets his or her first choice when voting honestly and can therefore do no better. For a voter who is not the median voter, any movement away from his or her first choice will have no effect if the vote that is cast stays on the same side of the median, and if the vote that is cast is on the other side of the median, then the action will move the chosen option one vote farther from the voter's actual first choice than the outcome that would have been selected with an honest vote. Since preferences are single-peaked by assumption, this can only make the voter worse off.

There is a reason for choosing randomly between the two medians in the case of an even number of voters, rather than picking the point midway between the two. If the point midway between the two medians were chosen, a voter who expected to be one of the median voters would have an incentive to vote for an outcome that was farther from the other median than his or her true first choice, so that the point midway between the two medians would move toward the voter's first choice. It is quite unlikely that anyone could know enough to employ such a strategy successfully when there were more than a very few voters, but the rule of choosing randomly between the two medians gives the rule the formal property of being completely resistant to strategizing.

The way that the UMC rule evades the Gibbard–Satterthwaite theorem is by not being a ranking-based rule. If one redefines the rule as a function of voters' rankings of the options, then either it violates the universal domain requirement, by limiting voters to rankings that are single-peaked, or it offers strategic opportunities to voters whose preferences are not single-peaked. Suppose for example that a voter's ranking of the options is such that, for any two options i and j, option i is ranked ahead of option j if and only if the height of the curve in Figure 14.1 is higher at i than at j.

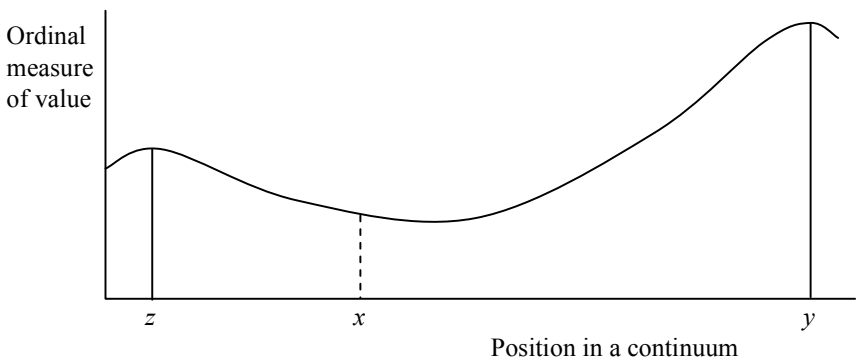


Figure 14.1 Strategic opportunities under the UMC rule

If the median is at x when the voter states honestly that his or her first choice is y, then falsely stating a first choice of z will move the outcome one vote to the left of x, which the voter will prefer.

Because of its simplicity, the median choice rule may be attractive when options lie in a one-dimensional continuum, even if there are some voters whose preferences are not single-peaked. This will be particularly true if there is reason to believe that the preponderance of voters have single-peaked preferences and that the general neighborhood of the outcome can be predicted from past experience, so that most of those whose preferences are not single-peaked will be able to see that they want to vote for a local peak of their preferences in the neighborhood of the outcome that will be selected. If there is reason to believe that a substantial proportion of voters have preferences that are not single-peaked, then it may be sensible to ignore the fact that options are in a continuum, allow voters to nominate options as discrete possibilities, and then use a vote-processing rule from Chapter 13.

One might think that it would be more natural to designate the mean, the average of the expressed first choices, rather than the median as the chosen option. The reason for designating the median is the very strong incentive that a mean-choice rule would create for strategic voting. A voter who believed that the mean of the choices of others would be less than what he or she wanted could raise the mean by voting for an outcome that was greater than what he or she actually wanted. The more the stated choice was raised, the higher the average would be. If a large number of voters responded to such incentives by trying to pull the outcome in the direction they believed advantageous, then there would be no reason to believe that the mean of the stated choices bore any relation to the mean of what people actually wanted.

Single-Stage Vote-Processing Rules for Continuums of Two or More Dimensions

As was mentioned in Chapter 9, when options lie in a continuum of two or more dimensions there is generally no majority-rule equilibrium. However, this does not preclude collective decisions. It simply means that instead of using open-ended sequences of choices by majority rule, collectivities must employ vote-processing rules that invariably terminate. Termination can be assured either by defining the outcome as a function of information obtained in a single stage of inquiry or by employing a multistage rule that explicitly excludes endless adjustment of the outcome. This section discusses four rules that operate in a single stage. The subsequent section discusses multistage versions of two of these.

The discussion of each rule introduced in this section includes classification in terms of two formal properties and three qualitative characteristics. The two formal properties are independence of rotations and independence of changes in

scale. The three qualitative characteristics are resistance to strategizing, ease of use, and ease of computing outcomes.

A vote-processing rule is independent of changes in scale if a change in the units in which one of the dimensions is measured, leaving the units in which other dimensions are measured unchanged, leaves the outcome unchanged. The UMC rule can be shown to possess this property. A change in scale does not change the location of the median voter or voters and therefore does not change the outcome.
A vote-processing rule is independent of rotations if a rotation of the coordinate system in which voters report their votes leaves the outcome unchanged. This property has no meaning for the UMC rule because the concept of rotation is not defined in one dimension.

The UMC rule motivates voters to state their preferences honestly (as long as their preferences are single-peaked), and therefore it is completely resistant to strategizing for single-peaked preferences. For multidimensional continuums, the natural generalization of single-peaked preferences is preferences that are "convex," that is, preferences such that a voter who is indifferent between two points prefers any point in the interior of the line segment connecting them to either of the end points. All four of the rules for multidimensional continuums that are discussed in this section offer voters opportunities for strategic voting even when preferences are convex. However, the degree of incentives to vote strategically, when preferences are convex, varies considerably among the rules. Therefore it is interesting to examine resistance to strategizing as a qualitative characteristic. Ease of use and ease of computing outcomes are straightforward qualitative characteristics.

The Multidimensional Median Choice Rule

One rule for choices in multidimensional continuums that comes immediately to mind is a generalization from the case of options in a one-dimensional continuum to the "multidimensional median choice" (MMC) rule:

> *Definition*: Multidimensional median choice rule: When a decision is to be made from a continuum of two or more dimensions, each voter shall state the value for which he or she wishes to vote in each dimension. From these statements, the median or medians in each dimension shall be determined. In each dimension in which the median is unique, that value shall be declared the winning value; in each dimension in which there are two medians that differ, a random process shall be used to choose between them.

The MMC rule is the simplest possible rule for multidimensional continuums in terms of ease of use and ease of computing outcomes. It is therefore given scores of 10 for both of these qualitative characteristics. Since decisions in the different

dimensions are made independently of one another, the MMC rule is independent of changes in scale.

The condition of single-peaked preferences, which suffices to motivate voters to vote for their actual first choices under the UMC rule, does not suffice to motivate honest voting under the MMC rule, because strategic opportunities arise under this rule when preferences in the different dimension are not independent of one another. These strategic opportunities can be illustrated as follows: A lack of independence means that what is most attractive in one dimension depends on what is chosen in another. Such preferences are expressed by indifference contours that are ellipses with axes that are not aligned with the dimensions of choice, as in Figure 14.2.

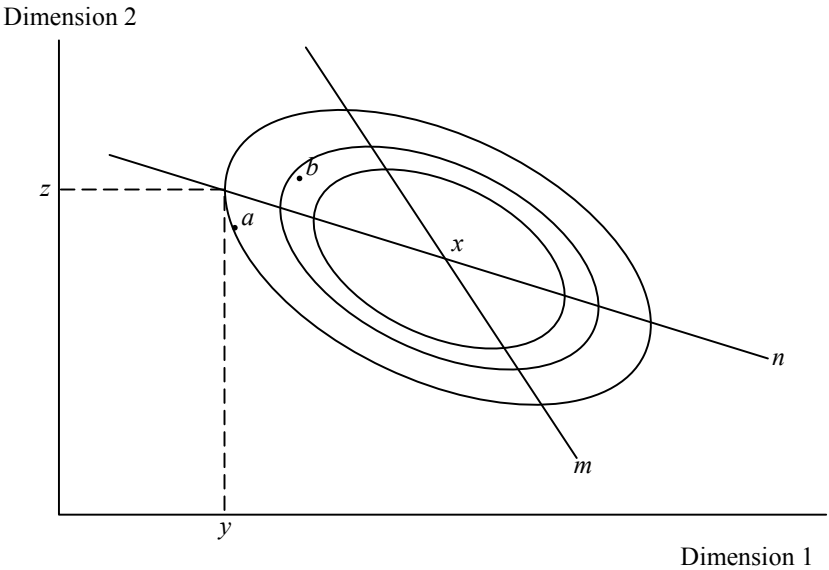


Figure 14.2 Non-independent preferences in two dimensions

Point *x* is the voter's first choice for the combination of the two dimensions, the voter's "ideal point." Each contour around point *x* is a set of points that are equally attractive to the voter, with all points inside any given contour more attractive than the points on or outside the contour. Note that if the level in Dimension 1 that is chosen is *y*, then the level in Dimension 2 that the voter finds most attractive is no longer described by *x*, but rather by *z*. Line *m* is composed of points that are most attractive in Dimension 1, given an outcome in Dimension 2, while line *n* is composed of points that are most attractive in Dimension 2, given an outcome in Dimension 1.

If the MMC rule is used, and if a voter whose preferences are given by Figure 14.2 expects an outcome in the neighborhood of *a*, then that voter has an incentive to name a point such as *b* as his or her choice, rather than the true choice *x*, so that the vote would pull the median in the direction that the voter preferred in the neighborhood of the expected outcome.

A sufficient condition for eliminating strategic opportunities under the MMC rule is that each voter's preferences in any one dimension are independent of what is chosen in any other dimension. Independent preferences generate indifference contours that have axes of symmetry that are aligned with the dimensions of choice. If every voter has such preferences, then what is most attractive in any one dimension is independent of what is chosen in any other dimension. The effect of a vote is always to pull the median in a direction in each dimension that the voter prefers. Independence of preferences in different dimensions is often an unreasonable assumption, and when it is violated the incentive for strategic misrepresentation of preferences can be quite strong. Therefore the MMC rule is given a score of 4 for resistance to strategizing.

The need for voters to engage in strategic thinking under the MMC rule, when the axes of indifference contours are not aligned with the axes of choice, suggests some variations on the MMC rule that might be attractive. If there was reason to believe that the preponderance of voters had preference contours with a given orientation, the choice might be redefined so that the axes of the choice corresponded to the axes of the preference ellipses. For example, if the choices were the amounts of money to be spent on schools and on road repairs, and if it was believed that the indifference contours of most voters were ellipses with axes at 45° from vertical and horizontal, then the choices could be redefined as the total amount of money to be spent and the difference between the amounts to be spent on schools and on roads. Such indifference contours would arise if people had strong feelings about the level of total spending that was appropriate and less strong feelings about how it ought to be allocated.

The idea of changing the axes of choices brings to light the fact that, under the MMC rule, the option that is selected varies with rotations of the axes of selection. This is illustrated in Figure 14.3.

Points *a*, *b* and *c* are the ideal points of three voters. When the choice is made in terms of Dimensions 1 and 2, point *x* is the outcome. At that point, the desire of voter *c* to move the outcome to the left is balanced by the desire of *b* to move it to the right, and *a* desires movement in neither direction, while in the vertical dimension the desire of *a* to move the outcome up is balanced by the desire of *b* to move it down, and *c* desires movement in neither direction. However, if the choice is made in terms of axes that are rotated 45° with respect to Dimensions 1 and 2, then point *y* is the outcome.

It is possible to devise a rule that eliminates almost all of the incentive for strategic responses that arises with the MMC rule. It is also possible to devise a rule that is independent of rotations of the axes. Finally, it is possible to devise a

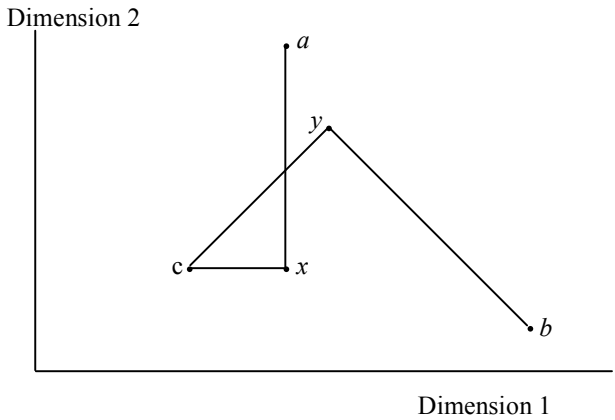


Figure 14.3 Outcomes that vary with rotations of axes under the MMC rule

rule that accomplishes both of these objective. These three rules are described in the next three sub-sections.

The Multidimensional Balanced Pulls Rule

The rule that eliminates almost all of the incentive for strategic responses is one that permits each voter to specify, at each point in the available continuum, the direction in which he or she would wish to pull the outcome in each dimension, if the outcome were there. Call this rule the "multidimensional balanced pulls" (MBP) rule:

> *Definition*: Multidimensional balanced pulls rule: When a decision is to be made from a continuum of two or more dimensions, each voter shall state the combination of values that he or she finds most attractive. Each voter shall also be given the chance to specify, for each dimension of choice, a "fixed set," that is, the (continuous) set of options, x, such that, if x were chosen, the voter would not wish to move the outcome in either direction in the specified dimension. For any voter who does not choose to make such specifications, it shall be assumed that the points in the fixed set in each dimension, d, all have the same value in dimension d as the voter's first choice. If the number of voters, n, is odd, then the "critical number" shall be defined as $(n + 1)/2$. If the number of voters is even, then a random process shall be used to decide, separately in each dimension, whether the critical number will be $n/2$ or $(n + 2)/2$. For each dimension, vote counters shall identify the "stable set," that is, the set of points such that the number of voters who are either satisfied with the value in the selected dimension or desire a smaller value, given the

values in the other dimensions, is equal to the critical number. The unique intersection of the stable sets shall be declared the winning option.

For the voter whose preferences are shown in Figure 14.2, the fixed set for Dimension 1 is line *m*, where indifference contours are horizontal, while for Dimension 2 it is line *n*, where indifference contours are vertical.

The operation of the MBP rule is illustrated in Figure 14.4. The choices of four voters, *a*, *b*, *c* and *d* are displayed. The line labeled a_1 is voter *a's* fixed set for dimension 1. The other lines are defined correspondingly. Since there are four voters, a random process must be used to select either 2 or 3 as the critical number in each dimension. If 2 is selected in both dimensions, then the outcome is *x*.

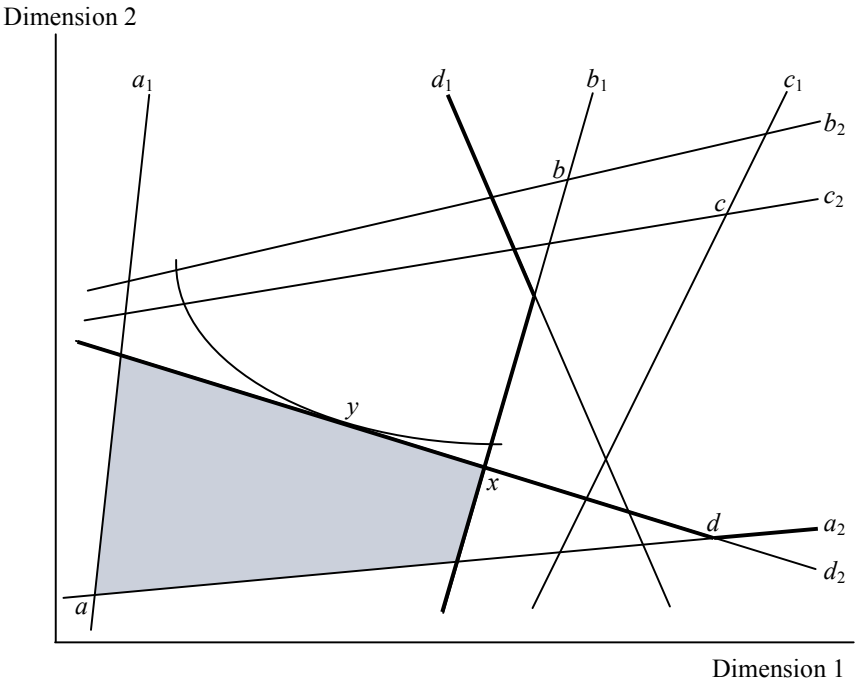


Figure 14.4 An illustration of the MBP rule

The reason that the MBP rule selects a unique outcome can be explained as follows: If a choice is to be made from a continuum of *k* dimensions, then each voter's fixed set for each dimension forms a continuous surface of $k-1$ dimensions. The stable set for each dimension will also be a continuous surface (e.g., the thick line in Figure 14.4). The intersection of the stable set for Dimension 1 with that for Dimension 2 is a surface of $k-2$ dimensions. The

intersection of this surface with the stable set for Dimension 3 is a surface of $k - 3$ dimensions, and so on until the intersection with the surface for Dimension k produces a surface of $k - k$ or zero dimensions, namely a point.

The MBP rule removes most but not quite all of the incentive to vote strategically. The remaining opportunity for strategic voting can be seen by considering the situation of voter *b* in Figure 14.4. If the critical number in both dimensions is 2, then the set of outcomes that are available to voter *b*, given the votes of the other voters, is represented by the shaded area. The point in this set that is most attractive to voter *b* is point *y*, where an indifference contour is tangent to the attainable set. However the point that results from an honest statement of *b's* preferences is *x*, where *b's* indifference contour is tangent to a horizontal line. Voter *b* succeeds in moving the outcome in the right general direction with an honest vote, but not to quite the most attractive point in the attainable set. However, to make a successful strategic vote, *b* would need to be able to predict the direction in which line segment *xy* would tilt. Furthermore, the benefit from successful strategic voting would be measured by the value to *b* of the relatively small difference between the contour on which *x* lies and the one on which *y* lies. On this basis, the MBP rule is given a score of 9 for resistance to strategizing.

The requirement that each voter name a continuous surface as his or her fixed set for each dimension ensures that the stable set for each dimension is continuous, and therefore that the stable sets intersect. If the fixed sets were not continuous, then, as with unidimensional preferences that are not single-peaked, there would be little reason for employing a spatial representation of the options. One might as well allow the voters to nominate options and use a vote-processing rule from Chapter 13.

The difficulty of using the MBP rule depends greatly on the number of dimensions. If there are just two dimensions, then voting requires only the drawing of two lines (or the specification of their equations). If there are three dimensions, then a voter must specify three surfaces. While it is possible that most voters could learn to use the MBP rule in two dimensions, it is questionable whether very many could learn to use it in three or more dimensions. Therefore it is given a score of 3 for ease of use. If the voters are able to express their preferences, the determination of the outcome is somewhat tedious but quite straightforward. Therefore the MBP rule is given a score of 8 for ease of computing outcomes.

The MBP rule is independent of changes in scale because changes in scale would not affect the locations of fixed sets. However, the MBP rule is not independent of rotations of the axes. This can be seen by considering the example shown in Figure 14.3 in terms of the MBP rule. If the three voters are assumed to have preference such that, between any two points, each voter prefers the point that is closer to his or her ideal point, then the MBP rule yields the same results as the MMC rule. Thus the change in the outcome for this example that occurs when the axes are rotated under the MMC rule would also occur under the MBP rule.

Two rules under which choices do not change with rotations of the axes will now be developed. The first of these has a rationale in terms of voters' preferences that depend only on distance from ideal points; the second has a rationale in terms of more general preferences.

The Minimum Sum of Distances Rule

If each voter's preferences depend only on distance from his or her ideal point, then a natural multidimensional generalization of the UMC rule is the minimum sum of distances (MSD) rule:

> *Definition*: Minimum sum of distances rule: When a decision is to be made from a continuum of two or more dimensions, each voter shall state the option for which he or she wishes to vote. The point that minimizes the sum of the distances to these points shall be declared the winning option. If the minimum is attained everywhere in an interval, then a random process shall be used to choose between the two end points of the interval.

While it is obvious that the MMC rule and the MBP rule reduce to the UMC rule for the case of a unidimensional decision, it is also true, but somewhat less obvious, that the MSD rule reduces to the UMC rule for unidimensional decisions. This follows from the fact that the point that minimizes the sum of the distances to a set of points in one dimension is the median of the set if the number of voters is odd, or any point in the closed interval between the two medians if the number of voters is even.

The MSD rule is independent of rotations of the axes of choice because rotation has no effect on distance measurements. For choices where it is possible to dispute what the axes ought to be, the MSD rule provides a natural generalization of the UMC rule to multiple dimensions while eliminating the need to determine the axes of choice.

While there is no analytic solution for the point that minimizes the sum of the distances to a set of points (it must be found by trial and error in successive iterations), there are some interesting ways of characterizing the MSD outcome for choices in two dimensions. Consider first the following analog device: On a stiff, flat surface such as a piece of plywood or a sheet of metal, set out a coordinate system describing the choice to be made. At each point named by a voter, drill a small hole and put a piece of thread through the hole, of a length equal to the longer of the dimensions of the surface. Tie all the threads together in a knot above the surface. Below the surface, attach uniform weights to the threads. Support the surface at a height high enough that none of the weights are on the ground. If the holes were of infinitesimal size, the threads of no weight, the knot of infinitesimal size, and there were no friction of the threads going through the holes, then the knot would come to rest at the point that minimized the sum of the

distances to the holes. The reason is that the device will come to rest in a configuration that minimizes its potential energy, subject to the constraints imposed by the placement of the holes. The minimization of potential energy will be characterized by a maximization of the sum of the lengths of thread below the surface. Since the total length of all threads is fixed, this will minimize the sum of the lengths above the surface. Thus the location of the knot is at the point that minimizes the sum of the distances to the holes. Where the MBP rule allows each voter to pull the outcome with the same force in each dimension of choice, the MSD rule gives each voter a single unit of force, which is allocated among dimensions according to the cosines of the angles between the axes and the direction in which the voter wishes to pull the outcome.

If there are three voters, and if the triangle formed by their ideal points has an angle greater than or equal to 120°, then the MSD choice is the vertex of that angle. If all angles of the triangle are less than 120°, then the MSD choice is the point such that the lines to the three chosen points form three angles of 120°. This is the point where the forces from the three weights sum to a zero vector. If there are four voters, and if one ideal point is inside the triangle formed by the other three, then the MSD choice is that point. Otherwise the MSD choice is the intersection of the diagonals of the quadrilateral formed by the four points. This point is again the point where the forces from the weights sum to a zero vector.

The MSD rule is not independent of changes in the units in which the axes are measured. For example, if the axes are dollars to be spent on schools and miles per hour for a speed limit, then a change in second axis to kilometers per hour could change the outcome. One easy way to see this is to note that if the axes are expanded or compressed, then the angle formed by the lines from a potential outcome to any pair of ideal points will change. Since such angles must be 120° in the case of three voters (as long as the vertex of the angle is the outcome and the triangle formed by the ideal points has no angle as great as 120°), changes in scale will generally change the outcome.

The MSD rule does not provide voters with incentives to report their preferences honestly, even when preferences are determined by distance. This is shown in Figure 14.5.

Points *a*, *b* and *c* represent the ideal points of three voters, placed in such a way that line segment *ab* has a length of $2\sqrt{3}$ and angle *abc* is 120°. For these conditions, the MSD rule selects point *b*. The possible outcomes that could result from different placements of point *c*, taking points *a* and *b* as given, form two arcs of circles of 120° each, with their end points at *a* and *b* and their centers at points *d* and *e*, which are midway between *a* and *b*, at a distance of 1 below and above line segment *ab*. (These arcs are the locus of vertices of angles of 120° between points *a* and *b*.) If the voter whose ideal point is at *c* orders possibilities according to their distances from his ideal point, then the point that is most attractive is not *b*, but rather *f*, which is where line *cd* intersects arc *aeb*. To achieve an outcome at *f*,

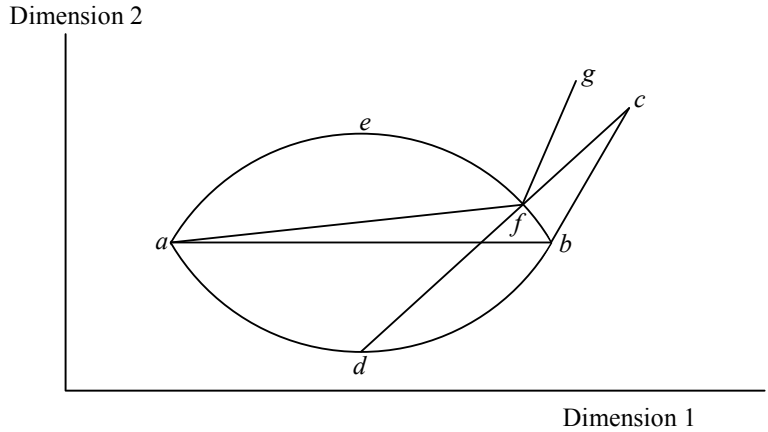


Figure 14.5 The incentive to misstate preferences under the MSD rule

the voter whose actual ideal point is *c* must state an ideal point along ray *fg*, which does not pass through *c* since angles *abc* and *afg* are both 120°.

In the case of preferences determined by distance, the lack of incentive compatibility is likely to be of little practical importance. The distance from *b* to *c* is only about 10 percent more than the distance from *f* to *c*. A more potent source of incentive to strategize under the MSD rule is the possibility that indifference contours would be ellipses rather than circles. As in Figure 14.2, voters who stated their first choices honestly could find that in some dimensions they were pulling the outcome in the opposite of the direction they desired. Since the source of incentives for strategic voting for the MSD rule is the same as for the MMC rule, the MSD rule is given a score of 4 for resistance to strategizing, the same as the MMC rule.

The MSD rule is as easy for voters to use as the MMC rule. Each voter needs only to state a preferred outcome. Therefore the MSD rule is given a score of 10 for ease of use. For vote counters, the rule would require computations that would only be practical on a computer. However, these are rather simple calculations. Therefore the MSD rule is given a score of 8 for ease of computation.

The Minimum Sum of Forces Rule

The MSD rule allows every voter to pull on the outcome with equal force and assumes that each voter wishes to pull in the direction of his or her ideal point. However, if the indifference contours of a voter are not concentric circles, then the voter will not generally want to pull the outcome in the direction of his or her ideal point. Therefore one might consider the "minimum sum of forces" (MSF) rule,

which is a variation on the MSD rule that allows voters to select the direction in which they will pull the outcome:

> *Definition*: Minimum sum of forces rule: When a decision is to be made from a continuum of two or more dimensions, each voter shall state the option for which he or she wishes to vote, and shall be given the opportunity to provide an indifference map of the available options. For a voter who does not provide an indifference map, it shall be assumed that indifference contours are concentric circles with centers at the voter's ideal point. At each point on each voter's indifference map other than the voter's ideal point, define the voter's "force" as a vector of length 1 orthogonal to an indifference contour in the direction of more preferred outcomes. At each point in the available set of possible outcomes, define the norm of the sum of the forces as the length of the vector obtained by summing of the forces of the voters whose ideal points are not at that point, minus the number of voters whose ideal points are at that point. The point that shall be selected as the outcome is the point where the norm of the sum of forces is minimized. If the minimum is attained everywhere in an interval, then a random process shall be used to choose between the two end points of the interval.

The instruction to incorporate the negative of the number of voters who have a point as their ideal point into the norm of the sum of the forces accommodates the fact that, at a voter's ideal point, the voter wishes to pull the outcome not in any one direction, but rather toward that point from any direction, canceling out one unit of force for any proposed movement from that point.

Like the MSD rule, the MSF rule is independent of rotations of the axes, because rotations do not affect the norm of the sum of forces. The MSF rule is not independent of changes in scale, because a change in scale does not preserve orthogonality, which is employed in the MSF rule.

For the case of voters who all have preferences determined by distance from their ideal points, the MSF rule reduces to the MSD rule, just as the MBP rule reduces to the MMC rule. Because the MSF rule reduces to the MSD rule for the case of preferences determined by distance, the proof that the MSD rule provides an incentive for strategic voting even when preferences are determined by distance proves that the MSF rule also provides incentives for strategic voting. However, by eliminating the unrealistic implicit assumption of preferences determined by distance, which is incorporated in the MSD rule, the MSF rule eliminates the preponderance of the incentive for misstating preferences. Therefore the MSF rule is given a score of 9 for resistance to strategizing.

It is likely to be quite challenging to devise a framework in which voters would be able to report their indifference maps, and it is likely to be difficult for voters to do so if they have the opportunity, particularly if there are more than two

dimensions in the continuum from which a choice must be made. Therefore the MSF rule is given a score of 2 for ease of use.

The problem of computing the outcome under the MSF rule involves first the difficulty of processing a collection of indifference maps and the relatively straightforward computational problem of finding the minimum of a function. Therefore the MSF rule is given a score of 6 for ease of computation.

Multistage Vote-Processing Rules for Continuums of Two or More Dimensions

The greater responsiveness to voters' preferences of the MBP and the MSF rules comes at a cost of a much reduced ease of use and ease of computing. Multistage vote-processing rules provide a way of achieving some of the advantages of these rules, with much greater ease of use and ease of computing outcomes. However, some resistance to strategizing is sacrificed.

The Multistage Multidimensional Balanced Pulls Rule

Consider first the "multistage multidimensional balanced pulls" (MMBP) rule:

> *Definition*: Multistage multidimensional balanced pulls rule: When a decision is to be made from a continuum of two or more dimensions, each voter shall state the option for which he or she wishes to vote. The option that is the median preference in all dimensions is identified and labeled the first provisional outcome. Each voter is then asked: "Taking the choices in all other dimensions as given by the first provisional outcome, what is your first choice in dimension x?" with x taking on successively the values of 1, 2, and so on through all the dimensions. The voters' responses are treated as if they described multidimensional first choices, and a new multidimensional median outcome is identified, and labeled the second provisional outcome. Voters are again asked to state a first choice in each dimension, now taking the choice in every other dimension as described by the second provisional outcome. The process iterates in this way until a stopping condition is attained. The stopping condition is that either two successive provisional outcomes are sufficiently close to each other that further refinement is not worthwhile, or that the current provisional outcome is farther from the previous provisional outcome than from an earlier provisional outcome, O_t. In this case the final choice is the average of the provisional outcomes back to O_t.

If voters' preferences in different dimensions are independent, then the first and second provisional outcomes will be equal, and the process will stop at that point. If voters always respond honestly, the process will converge to the MBP outcome.

However, if voters respond strategically, their responses in successive iterations may not be consistent with any coherent preferences. This can lead to a lack of convergence.

As with the single-stage MBP rule, the outcome of the MMBP rule generally varies with rotations of the outcome space, but is independent of changes in scale. The MMBP rule is given a score of 8 for resistance to strategizing, nearly as good as the MBP rule, since the process is likely to be nearly over before voters have enough information to do any useful strategizing. It is given a score of 8 for ease of use, since the questions that voters are asked to answer are no more difficult than with the MMC rule, but there are more of them, because of the multiple stages. It is given a score of 9 for ease of computing, since all that is needed is identification of means in individual dimensions.

The Multistage Minimum Sum of Forces Rule

The "multistage minimum sum of forces" (MMSF) rule is to the MMBP rule as the MSF rule is to the MBP rule.

> *Definition*: Multistage minimum sum of forces rule: When a decision is to be made from a continuum of two or more dimensions, each voter shall state the option for which he or she wishes to vote. The outcome that minimizes the sum of the distances to the stated first choices is identified and labeled the first provisional outcome. Then every voter is asked to specify a set movements away from the first provisional outcome that are equally unattractive, with each movement either in just one dimension or equal distances in two dimensions. These movements are used to infer a utility function for the voter, with concentric elliptical indifference contours centered on the voter's stated first choice. At each possible outcome, each voter is assumed to desire to pull the outcome in the direction orthogonal to his or her indifference contour, in the direction of his or her first choice. The second provisional outcome is the point such that the sum of such vectors of unit length for all voters is minimized. This point is labeled the second provisional outcome. Voters are asked to specify a set movements away from the second provisional outcome that are equally unattractive, with each movement in just one dimension. These movements are used to infer a new utility function for the voter, with concentric elliptical indifference contours centered on the voter's stated first choice. The process iterates in this way until a stopping condition is attained. The stopping condition is that either two successive provisional outcomes are sufficiently close to each other that further refinement is not worthwhile, or that the current provisional outcome is farther from the previous provisional outcome than from an earlier provisional outcome, O_t. In this case the final choice is the average of the provisional outcomes back to O_t.

The MMSF rule achieves independence of rotations at a cost of independence of changes in scale. The MMSF rule is given scores of 8 for resistance to strategizing and ease of use because it is like the MMBP rule in these regards. It is given a score of 7 for ease of computation because the computation of utility functions and points where pulling vectors sum to a zero vector is more complex than identifying medians. It is computationally simpler than the MSF rule because all indifference contours are standardized as concentric ellipses.

The MMSF rule is essentially the same as a mechanism proposed by Annund Hylland and Richard Zeckhauser.[1] Hylland and Zeckhauser suggested a decision-making mechanism in which voters would distribute "influence points" among issues, with the novel feature that each voter's influence on each issue would be equal to the square root of the number of points that the voter allocated to the issue. A search would be made for an equilibrium, where, for each issue, the influence expended on increasing the parameter of choice was equal to the influence expended on decreasing it.

Counting votes according to the square roots of the numbers of points assigned by voters guarantees that the sum of the squares of the points assigned to all issues by any voter is constant. The square root of this sum of squares can be considered the length of the vector with which the voter pulls the outcome. Thus the location where pulls in all directions are balanced is a point where the sum of pulling vectors is the zero vector.

Comparison of Rules for Choices from Continuums of Two or More Dimensions

Six rules for choices from continuums of two or more dimensions have been discussed: multidimensional median choice (MMC), multidimensional balanced pulls (MBP), minimum sum of distances (MSD), minimum sum of forces (MSF), multistage multidimensional balanced pulls (MMBP) and multistage minimum sum of forces (MMSF). The properties of these rules in terms of the five characteristics that have been examined are shown in Table 14.1.

One of the first things that might be noticed in Table 14.1 is that each rule has exactly one of the properties of independence of rotations and changes of scale. The three rules that operate on the various dimensions separately are independent of changes in scale but not of rotations, while the three rules that operate in all dimensions simultaneously are independent of rotations but not changes in scale. Next it might be noticed that two rules score high in ease of use and low in resistance to strategizing. These rules incorporate an implicit assumption of preferences determined by distance. They encourage strategizing when this assumption is not satisfied. Two other rules are high in resistance to strategizing but low in ease of use. The remaining two are moderately high in both ease of use and resistance to strategizing.

Table 14.1 Properties of six rules for selecting one option from a continuum

	Independence of Rotations	Independence of Changes in Scale	Resistance to Strategy	Ease of Use	Ease of Computing Outcomes
MMC	No	Yes	4	10	10
MBP	No	Yes	9	3	8
MSD	Yes	No	4	10	8
MSF	Yes	No	9	2	6
MMBP	No	Yes	8	8	9
MMSF	Yes*	No	8	8	7

* Assuming convergence

These patterns suggest that the choice of a vote-processing rule for a continuum of two or more dimensions can be made in the way shown in Table 14.2.

Table 14.2 The choice of a vote-processing rule for a continuum of two or more dimensions as a function of two criteria

		Importance of user simplicity		
		Low	Moderate	High
Obvious dimensional comparability	No	MBP	MMBP	MMC
	Yes	MSF	MMSF	MSD

When user simplicity is very important, one should choose either the MMC or the MSD rule; when it is not at all important, the MBP and MSF rules are attractive for their closer accord with what voters want and their resistance to strategizing. When user simplicity is moderately important, the MMBP and MMSF rules provide almost as much as the MBP and MSF rules while being only slightly more complex than the MMC and MSD rules.

If there is no obvious way of comparing a unit in one dimension with a unit in another, then there is no need for independence of rotations, and the MMC, MBP and MMBP rules suffice and provide independence of changes in scale. However, if there is an obvious way of comparing a unit in one dimension with a unit in another, then independence of changes in scale is not needed and the MSD, MSF and MMSF rules are attractive for their independence of rotations.

The next chapter discusses voting procedures for selecting more than one option.

Note

1 Hylland, Aanund and Richard Zeckhauser, "A Mechanism for Selecting Public Goods When Preferences Must be Elicited," Kennedy School of Government Discussion Paper 70D, Harvard University, August 1979.

Chapter 15

Vote Processing Rules For Selecting More than One Candidate When Votes Have Predetermined Weights: Proportional Representation

Introduction

This chapter is concerned with vote processing rules for selecting more than one candidate when candidates have predetermined weights and do not have identifiable locations in a continuum. The kind of vote-processing rule that is attractive in these circumstances depends on the goal of a collective decision. If the goal is to identify the objectively best set of candidates (as with disinterested judges selecting the members of a debating team) then it is sensible to have voters rank the candidates and either:

a. process the rankings by one of the rules discussed in Chapter 13 that generates an overall ranking (such as the ranked pairs rule) and select the candidates that are ranked highest in that overall ranking, or
b. process the rankings by one of the rules discussed in Chapter 13 that generates a top candidate (such as the maximin rule), select that candidate, delete all references to that candidate in the rankings, and then repeat the process of selecting the top candidate and deleting all references to that candidate until all positions have been filled.

Often, however, the goal of a collective decision is not to select the candidates that seem to be best but rather to select a set of candidates that reflects the diversity of opinions and/or interests in the electorate. It is for these situations that there are new considerations for the problem of devising an attractive vote-processing rule, which are the concern of this chapter.

It is generally assumed that the number of candidates to be selected has been fixed. However, Gordon Tullock has suggested that it may be attractive not to fix the number but rather to have a representative assembly in which each voter is represented by the person of his or her choice.[1] Each representative's vote would be weighted according to the number of voters represented. Speaking time in debates, salary, and other perquisites of office would be proportioned similarly. In this era of electronic interconnectedness, one can imagine including in such a

system the possibility of changing one's representative on a daily basis, just as one can change one's stockbroker. It is difficult to know whether the reason that such systems are not used is simply that people have not thought about them enough to be aware of their attractive properties, or whether there is some good reason why systems of representation ought to be of a more traditional form. Whatever the reason may be, the systems that collectivities actually use for selecting more than one candidate involve a fixed number of selected candidates. When the candidates are representatives, they all have the same voting power. This chapter is concerned with the evaluation of such systems.

The chapter begins with discussions of two relatively simple ways of selecting more than one candidate, namely multiple vote plurality and cumulative voting. The first is shown to be quite lacking in representativeness, while the second is only modestly representative. The chapter then moves to a discussion of party-list systems of proportional representation and issues in their construction. The remainder and greatest part of the chapter is concerned with ranking-based systems of proportional representation, which turn out to be variations of the single transferable vote (STV). Issues involved in the choice of a quota for STV and in determining which votes to transfer are discussed. Summaries of the rules of the most prominent versions of STV are then provided. These traditional versions of STV all use sequential eliminations and therefore lack Condorcet consistency and non-negative responsiveness when applied to the selection of one candidate. To avoid these difficulties, a new version of STV based on paired comparisons is developed. Principles for the evaluation of matrices of majorities based on such comparisons are discussed. These new ideas are then extended to the problems of casual vacancies and stratified representation. The chapter concludes with discussions of computational efficiency and the refinement–comprehensibility trade-off.

Multiple Vote Plurality

One way of selecting multiple candidates that is a natural extension of plurality is to permit every voter to vote for up to x candidates, for some chosen number x, and select the k candidates that receive the most votes. While this way of selecting multiple candidates is probably more common than any other, it has limitations that are similar to the limitations of plurality for selecting one candidate, as discussed in Chapter 13. If x is 1, there are in particular the problems of "vote splitting." Two or more candidates that appeal to the same constituency will split the vote of their supporters, and neither will be selected when either would have been selected if the other had not been on the ballot. The potential problem of vote splitting diminishes as x increases, but it does not disappear. Whatever value x takes, it would be possible that the number of candidates appealing to the same

constituency to be greater than x, so that the splitting of the votes among similar candidates would reduce the ability of the group to have their candidates selected.

If x is equal to k or more, then there is an additional criticism that can be made against multiple vote plurality. A faction that constitutes more than half of the electorate can organize their supporters to vote for the same k candidates and be assured that they will secure all k positions. The minority will be entirely unrepresented. If it is intended that the selected candidates *represent* the electorate, then such a system does not succeed. There is no principle that gives any particular value of x a claim to being the right value.

Cumulative Voting

One way of avoiding some of the worst problems of multiple vote plurality is to give each voter k votes that can be allocated among the candidates as the voter chooses. The k candidates with the most votes are selected. Such a system is called cumulative voting. Cumulative voting is often used for selecting corporate boards of directors.

Let n denote the number of voters. Under cumulative voting, any group of more than $n/(k+1)$ voters can secure the selection of the candidate they favor by allocating all of their votes to their chosen candidate. There can be at most k other candidates with as many votes, so their candidate must be among the winners. Thus cumulative voting provides an important element of proportional representation. However, cumulative voting requires voters seeking representation of their views to estimate their support accurately if they are to be represented proportionally. If a group splits their votes among two candidates because they believe they have enough support to secure the selection of two, they can wind up with neither of their candidates selected if they do not have as much support as they expected, when they could have had one of their candidates selected by giving that candidate all of their votes. Similarly, a group may underestimate its support and achieve the selection of only one candidate when they had enough votes to have two selected if they had divided their votes between the two.

Party-List Systems

In party-list systems, voters vote for parties, and representatives are allocated to parties according to the number of votes that parties receive. Position on lists drawn up by the parties is the primary determinant of which persons from each party are selected, although European systems of proportional representation of the party-list type all have added features to give voters some voice in the selection of representatives within parties.[2]

There are three principal issues in party-list systems:

1. How will the representatives from each party be selected?
2. How will fractional seats be apportioned?
3. Will there be some threshold, greater than the number of votes otherwise needed to secure one seat, below which a party will be awarded no seats?

In a pure party-list system of proportional representation, such as that used in Israel, the only information that a voter provides is the party for which he or she chooses to vote. Before the election each party provides a list of its candidates. The parties are allocated positions in proportion to the numbers of votes they receive, and if a party obtains x seats, the first x candidates on its list are selected.

Recognizing that it may be unattractive to give party officials sole power to determine the order in which a party's candidates are chosen, some systems of proportional representation give a voter the power to name a candidate within the party for which he or she votes. Any candidate that receives n/k votes is selected, and the party list is used only to fill the seats that the party is awarded but does not have candidates with enough personal votes to fill. While such a system gives voters noticeably more power than pure party-list systems, they have the disadvantage that a faction within a party that wants its own candidates selected must calculate its support accurately and assign different supporters to vote for different candidates. The excess support of one candidate cannot be transferred to another.

The problem of allocating fractions of positions among parties in party-list systems of proportional representation is equivalent to the problem of allocating fractions of representatives among political subdivisions (such as states) in a federal legislative body in which representation is proportional to population (such as the U.S. House of Representatives). A variety of proposals for dealing with the problem have been advanced over the years.[3] After reviewing these proposals, Balinski and Young conclude that the most representative way of dealing with fractions is the way first proposed by Daniel Webster in 1832 for allocating representatives in response to the census of 1830.[4] With the size of the body to be selected having been decided, the method is to then find a divisor such that when the number of votes of each party (or population to be represented in the case of a body such as the U.S. House of Representatives) is divided by the divisor and rounded to the nearest whole number, the chosen number of representatives is awarded.[5] Webster's method, when invented independently by others who were concerned with the problem of proportional representation, has been called the method of major fractions, Sante-Lagüe's method, and the method of odd numbers.[6]

Under Webster's method, there are some circumstances in which a party can increase its representation by dividing into two parties. A party that has enough votes to receive 1 1/3 positions, which will be rounded down to 1, might be able to split into two parties each of which would be entitled to 2/3 of a seat, thereby receiving 2 seats when these numbers were rounded up. If this incentive is to be

avoided, then the appropriate method is the one proposed by Thomas Jefferson for apportioning seats in response to the first U.S. census in 1790. This is like Webster's method, except that all fractions are rounded down instead of to the nearest whole number.[7] Jefferson's method is also known as the method of greatest divisors, d'Hondt's method, the Hagenbach–Bischoff method, and the method of highest averages.[8]

Several countries that have party-list systems of proportional representation impose an additional rule that a party must obtain some number of votes greater than the number that is otherwise required, in order to receive any seats. The stated rationale for such a rule is that otherwise the selected group will be "excessively fractured." It is hard to know whether this is a reasonable concern or a device of the major parties that have power to make it more difficult for newcomers to challenge them.

Ranking-Based Systems

With a ranking-based system for selecting multiple candidates, a vote is a ranking of some or all of the candidates, ranked by attractiveness to the voter. Unlike party-list systems, there is no presumption that voters desire to be represented by parties and no presumption that a voter would rank all the candidates of one party ahead of all candidates of other parties. The set of candidates who are elected depends in a complex way on the rankings that are submitted. When the number of candidates to be elected is very large, the electorate may be divided into multi-member constituencies, so that voters are not expected to provide rankings of an unreasonably large number of candidates, and so that a greater degree of individual representation can be achieved. The most prominent ranking-based system for selecting multiple candidates is the single transferable vote (STV). STV is sometimes called the Hare system of proportional representation, although Thomas Hare was neither the first nor the last of many who contributed to the development of STV.[9]

STV is the predominant form of proportional representation of countries with a heritage of English influence. It is used to elect the Dàil (Assembly) in Ireland, the Senate in Australia, the House of Representatives in Malta, and some local bodies in New Zealand. In the U.S., STV is used to elect the City Council and School Committee in Cambridge, Massachusetts. It is used as well by dozens of unions and religious, charitable, and professional organizations in many countries. STV is not actually a specific vote processing rule, but rather a family of rules that share certain features and goals.

The earliest known suggestion that legislative bodies should represent the electorate proportionally occurred in a speech by Mirabeau to the Assembly of Provence, in 1789, but no actual mechanism was proposed. Saint-Just proposed in 1793 that the French Parliament be elected by a vote over the whole country as one

district, with each voter having one vote, and the candidates with the most votes winning. While this system, sometimes known as the single non-transferable vote, embodies the spirit of proportional representation, it suffers from the likelihood that some voters will be underrepresented when the candidates they vote for receive many times as many votes as they need to be elected, while other voters can be unrepresented because they vote for candidates who do not secure enough votes to be elected.

This criticism of the single non-transferable vote illustrates the need for a logically precise specification of proportionality in representation. The concept of "proportionality for solid coalitions" or PSC, developed by Michael Dummett, is useful for this purpose.[10] Dummett defines the set of voters V to be "solidly committed" to the set of candidates C if every voter in V ranks every candidate in C ahead of every candidate that is not in C. If a set of solidly committed voters is large enough, one or more of the candidates in the set to whom it is solidly committed should be elected. But how large is "large enough?"

The most obvious criterion would be that if there are n voters and k positions to be filled, and the voters in V are solidly committed to the candidates in C, then someone from C should be elected if the number of voters in V is greater than or equal to n/k. (The symbol $|V|$ will be used hereafter to denote the number of elements in set V. The number such that voters of that number must be represented when they are solidly committed to a set of candidates, e.g., n/k, will be called the "quota.") Setting the quota at n/k means, for example, that if two positions are to be filled, a solid coalition consisting of half or more of the electorate must be represented. This idea can be extended to multiple representation for a solid coalition by noting that if a solid coalition contains j quotas of voters who are solidly committed to the set of candidates C, for some positive integer j, and if there are j or more candidates in C, then j candidates from C ought to be elected. Formally, a voting rule possesses k-proportionality for solid coalitions, or k-PSC, if:

When there are n voters and k positions to be filled, then for any set of voters, V, and any set of candidates, C, and any positive integer, j,

$$\begin{bmatrix} V \text{ is a solid coalition} \\ \text{for } C \text{ and } |V| \geq jn/k \end{bmatrix} \Rightarrow \begin{bmatrix} \text{if } |C| \leq j, \text{ then all elements of } C \text{ are elected;} \\ \text{otherwise at least } j \text{ elements of } C \text{ are elected} \end{bmatrix}$$

The antecedent of k-PSC, however, may require too much of solid coalitions. With two positions to be filled, a solid coalition of 34 percent of the voters can rightly say that there could be at most one other solid coalition with a size greater than or equal to theirs, and on that basis they deserved one of the two positions. More generally, the solid coalition V should be represented if $|V| > n/(k+1)$. Incorporating the possibility that a solid coalition could be so numerous that it

would deserve to be represented by more than one candidate provides the definition of $(k+1)$-proportionality for solid coalitions, or $(k+1)$-PSC. A voting rule possesses $(k+1)$-PSC if:

When there are n voters and k positions to be filled, then for any set of voters, V, and any set of candidates, C, and any positive integer, j,

$$\begin{bmatrix} V \text{ is a solid coalition for} \\ C \text{ and } |V| \geq jn/(\mathrm{k}+1) \end{bmatrix} \Rightarrow \begin{bmatrix} \text{if } |C| \leq j, \text{ then all elements of } C \text{ are elected;} \\ \text{otherwise at least } j \text{ elements of } C \text{ are elected} \end{bmatrix}$$

For $k = 1$, $j = 1$ and $|C| = 1$, $(k+1)$-PSC reduces to the majority rule principle. That is, $(k+1)$-PSC says under these conditions that when just one candidate is to be selected, if more than half of the electorate ranks some candidate, A, ahead of all other candidates, then A is selected. By contrast, k-PSC reduces to the Pareto principle when $k = 1$, $j = 1$ and $|C| = 1$. That is, k-PSC says under these conditions that when just one candidate is to be selected, if *all* of the electorate ranks A ahead of all other candidates, then A is selected.

The earliest proposal to achieve the spirit of k-PSC was the rule that Thomas Wright Hill developed in 1819, for electing a committee of The Birmingham (England) Society for Literary and Scientific Improvement. Hill's rule involved iterative voting rather than the solicitation of voters' rankings of the candidates. Voters wrote their names on their ballots, and each voted for a single representative. Any candidate who received five or more votes was elected. If any candidate received more than five votes, ballots equal in number to the excess above five votes were selected at random, and the voters who had submitted them were invited to vote for alternative representatives. After iterating until there were no further surplus votes, the voters whose votes had elected no one were invited to vote again, until there were fewer than five votes that had not been used to elect anyone.

While serving as Secretary of the Colonization Commission of South Australia in 1839, Thomas Hill's son Rowland Hill arranged for an Australian municipal election to be held by a procedure inspired by his father's ideas, though not as detailed. A schedule of columns headed by the names of candidates was established, each voter was permitted to write his name in one column, and any candidate who achieved a specified number of names in his column was elected.

The sense in which Thomas Hill's rule satisfies the spirit of k-PSC is as follows: The quota of 5 for Hill's rule is the integer part of n/k. If there is a group of 5 voters who are solidly committed to the set of candidates C (though not all ranking the same candidate first), then in most cases (though it cannot be guaranteed) one would expect the group to succeed eventually in all assigning their votes to the same candidate, who would then be elected. If the number of voters in the solid coalition is some multiple of 5 and the coalition gives more than 5 votes

to some candidate, there will be at most 5 votes from the coalition used to elect this candidate (there could be votes for this candidate from outside the solid coalition), and the surplus votes can be used to elect other candidates that the coalition favors, so that the coalition should eventually succeed in electing at least one of the candidates it favors for every 5 members of the coalition.

The idea of having voters rank the candidates, instead of voting iteratively, was first put forward in 1856 by Carl Andrae, a Danish mathematician and politician. Without knowing about Hill's work, Andrae had the idea of establishing a quota of n/k (rounded down to an integer), examining the ballots in a random order, and if the candidate at the top of a voter's ranking was already elected, the vote would go to the first unelected candidate in the voter's ranking. Andrae's rule did not include, as Hill's had, the possibility of reallocating the votes of losing candidates. Andrae arranged for some Danish elections to be held using the rule he proposed, but the use of his rule soon died out.

The person who gave the idea of ranking-based proportional representation the publicity it needed to become a permanent part of political thinking was the London barrister, Thomas Hare. Hare thought the idea was original with him, and published his first book on the subject in 1859. His initial proposal was quite similar to Andrae's, but in 1865 he incorporated in his proposal the provision that after all surplus votes had been transferred, if the prescribed number of positions had not been filled, candidates would be eliminated successively, the one with the fewest votes first, and as each candidate was eliminated, each of his votes would be transferred to the unelected and uneliminated candidate ranked highest by the voter who supplied that vote. Hare thus recovered an important virtue of Hill's proposal.

The Choice of a Quota

The next significant improvement came in 1868, when another London barrister, H. R. Droop, proposed that the quota be reduced from the integer part of n/k to the integer part of $[n/(k+1)]+1$. The importance of this reduction in the quota is shown by the following example. Suppose that two Democrats, *D* and *E*, and two Republicans, *R* and *S*, are running in an election for three positions. There are 100 votes, distributed as follows:

24	23	32	21
D	*E*	*R*	*S*
E	*D*	*S*	*R*
R	*S*	*D*	*E*
S	*R*	*E*	*D*

Since a majority of the electorate is solidly committed to the two Republicans, one would expect the Republicans to be awarded two of the three positions. But if the quota is set at 33 (one-third of the electorate, rounded down to an integer) then no candidate receives a quota of first-place votes, and the candidate with the fewest votes, namely *S*, is eliminated, so that Democrats are awarded two of the three seats even though only 47 percent of the electorate favor the Democrats. On the other hand, if Droop's suggestion is employed, then the quota is 26. *R* is elected with a surplus of 6, which is transferred to *S*, securing *S*'s election, and the faction with a majority of the vote is awarded the majority of the seats. In view of examples of this sort, it is relatively non-controversial today that in establishing the quota, the number of voters should be divided by one more than the number of positions, rather than by the number of positions, as Andrae and Hare had suggested.

The introduction of the Droop quota improves the performance of STV in achieving proportionality in representation, but for some examples it still does not achieve the desired result. Suppose, for example, that four Democrats, *D*, *E*, *F*, and *G*, and four Republicans, *R*, *S*, *T*, and *U*, are contesting an election in which there are 800 voters and seven positions to be filled. The votes are distributed as follows:

100	100	100	99	101	101	101	98
D	*E*	*F*	*G*	*R*	*S*	*T*	*U*
E	*F*	*G*	*D*	*S*	*T*	*U*	*R*
F	*G*	*D*	*E*	*T*	*U*	*R*	*S*
G	*D*	*E*	*F*	*U*	*R*	*S*	*T*
R	*S*	*T*	*U*	*D*	*E*	*F*	*G*
S	*T*	*U*	*R*	*E*	*F*	*G*	*D*
T	*U*	*R*	*S*	*F*	*G*	*D*	*E*
U	*R*	*S*	*T*	*G*	*D*	*E*	*F*

Since 399 voters are solidly committed to the Democrats and 401 to the Republicans, one would expect the majority of the positions to go to the Republicans. However, the Droop quota is $[800/(7 + 1)] + 1 = 101$. Therefore the three Republicans with 101 votes are elected, and then the fourth Republican is eliminated, resulting in the election of the four Democrats.

What is needed to achieve the desired result is a quota of $n/(k+1)$. This is the quota employed in rules devised by Robert Newland and Frank Britton.[11] Incorporating a quota of $n/(k+1)$ into Hare's proposal provides a rule that satisfies $(k+1)$-PSC. The proof of this is as follows: Call $n/(k+1)$ an NB quota. Suppose that a coalition with a size greater than j NB quotas is solidly committed to the set of candidates C, with $|C| \geq j$. Then the number of voters *not* in the coalition is less than $n(k+1-j)/(k+1)$, which is not enough to provide $(k+1-j)$ NB quotas, though it may be enough to provide $(k-j)$ NB quotas. Therefore the number of

candidates that can be elected without using votes from the coalition is at most $k-j$. The votes of voters in the coalition are assigned to elements of C, and since the coalition is solidly committed to C, they remain assigned to elements of C after every transfer of a surplus or transfer of votes following the elimination of a candidate, as long as there are elements of C that remain unelected and uneliminated. Since the coalition has enough members to elect j candidates, it is not possible for the last element of C to be eliminated until j elements of C have been elected. And since those outside the coalition can elect at most $k-j$ candidates without votes from the coalition, the election cannot end before j candidates from C have been elected. Dummett offers a similar proof.[12]

While this is a strong argument in favor of the NB quota, it is possible to object to the NB quota on the ground that ties are not handled as well as they might be. For example, suppose that four candidates are competing for two positions and the votes are distributed as follows:

4	4	3	1
D	*E*	*F*	*G*
F	*F*	*G*	*F*
G	*G*	*E*	*D*
E	*D*	*D*	*E*

Under the NB quota, candidates D and E are elected, where a more even-handed procedure would note that there is a coalition of four voters solidly committed to the set $\{F, G\}$, and, since F is the choice of the majority of this coalition, the procedure would eliminate G, and then eliminate one of D, E and F with probabilities of 1/3 each. The mechanism for achieving this has been offered by Irwin Mann. In a computer program he wrote in 1973, Mann used the NB quota, with the added provision that a candidate was not declared elected until he or she had *more* than a quota of votes. Thus, for this example no one would be elected upon the distribution of the first-place votes, G would be eliminated, his vote transferred to F, and in view of the resulting tie, one of D, E, and F would be chosen by lot to be eliminated.

Determining Which Votes to Transfer

All of the STV systems through and including Hare's and Droop's entailed a stochastic element in the choice of which votes to transfer, and therefore a stochastic element in the final outcome. One might expect an attractive rule to yield the same outcome every time a set of votes was counted, apart from the possibility of true ties.

One step toward reducing the stochastic element in the outcome, used in Ireland, is to transfer not a random sample of all of the votes of a candidate with a

surplus, but rather a stratified random sample, after sorting the ballots according to who is named next on them. But because these ballots might be transferred again, a stochastic component remains.

The first version of STV to eliminate completely the stochastic component in the outcome, apart from true ties, was the "Senatorial Rules" developed in about 1880 by J. B. Gregory of Tasmania, Australia. The Senatorial rules treat each voter as having 100 votes. When there is a surplus, instead of selecting ballots at random to be transferred, the same number of a voter's 100 votes is transferred for each of the voters whose votes are in the "bundle" that generated the surplus. The same effect that Gregory achieved by treating each voter as having 100 votes is achieved in subsequent rules by transferring fractions of a vote.

There is no controversy about the way the idea of fractional transfers of a voter's vote ought to be applied to the transfer of a single surplus. However, for the case of more than one transfer there is a sharp difference of opinion about what procedure is appropriate. Consider the following example of 80 voters and five candidates competing for three positions:

25	15	9	9	6	7	9
R	*R*	*S*	*S*	*T*	*U*	*V*
S	*S*	*R*	*U*	*U*	*T*	*T*
V	*T*	*T*	*T*	*R*	*S*	*U*
U	*U*	*U*	*R*	*S*	*R*	*R*
T	*V*	*V*	*V*	*V*	*V*	*S*

While the actual Senatorial rules employed the Droop quota, I will use the NB quota here, to facilitate comparisons with other rules that use the NB quota.

Under the Senatorial rules, the 80 voters have a total of 8,000 votes. Thus the NB quota is 8,000/(3 + 1) = 2,000. Candidate *R* has 4,000 first-place votes all of which name *S* as the second choice. Since *R* needs only 2,000 of these to be elected, each of the voters who voted for *R* has 50 of his 100 votes is transferred to *S*, which when combined with *S*'s 1,800 first-place votes, give *S* 3,800 votes, so that *S* is elected also. Under the Senatorial rules, each of the 40 voters who had 50 votes transferred to *S* now has the 45 of those votes that *S* does not need transferred to his third choice, which is *V* for 25 voters and *T* for 15 voters. The 1,125 votes thereby transferred to *V* raise *V*'s vote total to 2,025, and *V* secures the third position.

The practice in the Senatorial rules, when transferring a surplus, of transferring only votes from the "bundle" that created the surplus (the votes that went for *R* to *S* in this case) is followed in all other rules I know of for counting STV votes by hand. While this may be tolerable if the extra cost of a more sophisticated procedure is too great, one can object to the practice on the ground that some of the power to elect candidates other than *R* and *S* should be given to the voters who named *S* as their first choice.

The first system for allowing all votes for a candidate with a surplus to participate in the distribution of the surplus was published by Brian Meek.[13] Meek's proposal was that, for each candidate with a surplus, a fraction between 0 and 1 be calculated, such that when the candidate retains that fraction of every vote or fractional vote that comes his way, the candidate winds up with exactly a quota. Applying Meek's rule to the example above, the fractions for *R* and *S* (call them *r* and *s* respectively) must satisfy the equations,

$$40r + 9(1 - s)r = 20;$$
$$18s + 40(1 - r)s = 20.$$

With a bit of algebra, the solution to this pair of equations can be derived to be (approximately)

$$r = .4494;$$
$$s = .4997.$$

The implication of these numbers is that each of the 25 votes with ranking *RSVUT* is transferred to *V* as (1 – .4494)(1 – .4997), or .2754 of a vote, each of the 15 votes with ranking *RSTUV* is transferred to *T* as .2754 of a vote, each of the 9 votes with ranking *SRTUV* is transferred to *T* as .2754 of a vote, and each of the 9 votes with ranking *SUTRV* is transferred to *U* as 1 – .4997, or .5003 of a vote. After these transfers, the vote totals (upon calculation with greater precision) are:

R 20.000;
S 20.000;
T 12.611;
U 11.503;
V 15.886.

Since there are no further surpluses to be transferred, the candidate with the fewest votes, *U*, is eliminated, and all of his votes are transferred to *T*, which gives *T* the position that went to *V* under the senatorial rules.

If there are more than two surpluses, one would almost certainly want to use a computer and a method of successive approximations to implement the Meek rule, as the degree of the equations is equal to the number of surpluses. On the other hand, once a solution has been derived, it is quite simple to confirm that it is the solution. A computer program implementing the Meek rule has been published by Hill, Wichmann and Woodall,[14] and the Meek method has been adopted by Royal Statistical Society for their elections.

The Meek method has been criticized by C.H.E. Warren for the way that it allocates the "cost" of electing a candidate among the candidate's supporters.[15] Under the Meek method, a candidate's quota is acquired by appropriating the same

proportion from every whole or fractional vote that the candidate receives. Thus in the example above, *S*'s quota is achieved by using 49.97 percent of each of the 18 votes that rank *S* first (a total of 8.99 votes), and 49.97 percent of each of the 40 votes of weight .5506 that are transferred from *R* to *S* (that is, 0.2751 from each of the 40 votes for a total of 11.01 votes).

Warren argues that it is inappropriate for the voters whose rankings begin *RS* to be "charged" less (0.2751 of a vote per person) for the election of *S* than the voters who ranked *S* first (0.4997 of a vote per person). Warren suggests that instead of allocating the same fraction, f_i, of each whole or fractional vote for the election of a given candidate, the counting rule ought to charge every supporter of a candidate who is elected the same price (percentage of a whole vote), p_i, provided that there is enough of the vote left to pay the price, irrespective of the fraction of the vote that is left when the vote reaches the candidate who needs part of it. Thus in the example above, since there are 49 supporters of *R*, each would be charged 20/49, or 40.82 percent, of a vote toward the election of *R*, while each of 58 supporters of *S* would be charged 20/58, or 34.48 percent, of a vote to the election of *S*. Then each of the 25 votes with ranking *RSVUT* is transferred to *V* as (1 – .4082 – .3448), or .2470 of a vote, each of the 15 votes with ranking *RSTUV* is transferred to *T* as .2470 of a vote, each of the 9 votes with ranking *SRTUV* is transferred to *T* as .2470 of a vote, and each of the 9 the votes with ranking *SUTRV* is transferred to *U* as 1 – .3448, or .6552 of a vote. After these transfers, the vote totals (upon calculation with greater precision) are:

R 20.000;
S 20.000;
T 11.928;
U 12.897;
V 15.175.

Thus under the Warren rule, candidate *T* is eliminated, his votes are transferred to *U*, and *U* is awarded the third position. I agree with Warren's views on this issue.

A second controversy with respect to the way a surplus is distributed arises with respect to the treatment of votes that cannot be transferred because they provide only partial rankings. The NB rules provide (as do other hand-counting rules) that when votes that cannot be transferred are encountered in the transfer of a surplus (of some candidate *A*), the weight of the votes that *can* be transferred is increased to offset the non-transferable votes, but not above what their weight had been as votes for *A*. One effect of this rule is to limit the fall in the total number of votes allocated to all candidates. This number falls in any case when the votes of an eliminated candidate cannot be transferred, but keeping the number higher reduces the extent to which, at the end of an election, there are candidates who are elected by default with much less than a quota because all other candidates have been eliminated. Thus the incorporation of this provision into the NB rule is

understandable. However, I have heard the suggestion that Meek's method should use a similar rule, rather than compensate for the non-transferability of votes by reducing the quota.

I see this controversy as follows: When a voter submits an incomplete ranking (a ranking that omits more than one candidate), he is "partially" or "conditionally" abstaining. He is saying, in effect, "If there comes a point in the count at which it is not possible for all or part of my vote to be allocated to any of the candidates I have ranked, then I desire that the remaining part be allocated to no one." The fact that one voter declines to use some of the voting power that is available to him means that some other voters must be accorded additional relative power. The practice of increasing the weight of transferable votes gives this additional relative power to the other voters whose votes were in the same bundle as that of the partial abstainer, while the Meek procedure, by lowering the quota, accords this additional relative power to all voters. If the goal of the election procedure were to devise a statistical estimate of how partial abstainers would have voted if they had chosen to complete their votes, then some variation on the NB procedure would be appropriate. However, in view of the fact that it is not general practice in elections to search for a proxy for each voter who abstains, it seems to me that the Meek procedure of reducing the quota, and thereby giving the additional power to all voters, is the more appropriate one.

Summary of STV Rules

Now that most of the issues surrounding STV have been introduced, it is time to provide more detailed specifications of the principal STV algorithms: the hand-counting rules of the Electoral Reform Society, the Meek rules and the Warren rules.

The hand-counting rules come from the booklet, "How to Conduct an Election by the Single Transferable Vote," published by the Electoral Reform Society of Great Britain and Ireland (hereafter the ERS). The ERS is the principal advocate for STV in the world today. It was formed in London in 1884 as The Proportional Representation Society. When its initial efforts were unsuccessful it fell moribund for about 20 years, but since 1906 it has been continuously active in the promotion of STV. In 1959 the Society's present name was adopted, to avoid the possibility that people would think that it promoted party-list systems of proportional representation.

The ERS rules for counting an election provide an algorithm for proceeding from any set of voters' rankings of candidates to a set of elected candidates of a specified size. (Because of the possibility of ties that are resolved by random processes, this and other STV algorithms cannot be described as functions.) Ties in voters' rankings are not allowed, but partial rankings (of only some of the candidates) are allowed. The vote counting begins by computing the quota,

defined as $n/(k+1)$, rounded up to the nearest hundredth. The quota will be designated hereafter by Q. Each vote is assigned to the candidate ranked first on it. Any candidate who receives Q votes or more is elected.

If one or more candidates received more than Q votes, these "surpluses" (excesses of the number of votes received above Q) are transferred sequentially to other candidates, as fractions of votes, beginning with the largest surplus. If a candidate has $S + Q$ votes, then each of these is transferred as the fraction $S/(S + Q)$ of a vote, to the "active" (not yet elected and not yet eliminated) candidate ranked highest by that voter.

After all surpluses have been transferred, as long as k candidates have not yet been selected, the candidate with the fewest votes is eliminated, and each vote for the eliminated candidate is transferred to the active candidate ranked highest on that ballot, at whatever fractional or full value the vote has. If these transfers push one or more candidates up to or above the quota, then those candidates are elected, and their surpluses are distributed. The process of declaring any candidate who attains the quota elected, transferring surpluses, and eliminating the candidate with the fewest votes continues until k candidates are elected. In the event of a tie among surpluses to be distributed or candidates to be eliminated, candidates are ranked according to votes possessed in the initial and then subsequent stages of the count. Only if all such comparisons yield ties is a random process used to break the tie.

Three other features of the rules recommended by the ERS should be mentioned. First, if the total of the untransferred surpluses is so small that it is impossible for them to affect the identity of the candidate to be eliminated next, then the transfer of the surplus, which can be a tedious process, is postponed until it can make a difference. This means that surpluses may be left untransferred at the end of an election. Second, if there is a set of candidates that cannot avoid elimination no matter how prospective transfers of surpluses turn out, then they are eliminated together rather than sequentially. Third, when some of the votes for a candidate with a surplus are non-transferable (because some voters have ranked only some of the candidates), the fraction of the transferable votes that is transferred is increased: the fraction transferred is the ratio of the surplus to the number of transferable votes, but not more than 1.[16]

Under the rules proposed by Meek, the initial quota is $n/(k+1)$, to whatever accuracy a computer computes the fraction. As with the ERS rules, votes are allocated to the candidates ranked first on them, and any candidate with a quota of votes is elected. Surpluses are then distributed simultaneously, by finding a set of fractions for elected candidates f_i, such that when each elected candidate i keeps the fraction f_i of the votes that come to it and passes the remainder to the candidate ranked next on that vote, every elected candidate has exactly Q votes. If some votes are untransferable because voters have not ranked all of the candidates, Q is reduced to the sum of the votes of all candidates divided by $(k+1)$. Unelected candidates keep all of the votes that come to them. If any unelected candidate

acquires Q or more votes in this process, its status is changed to elected, and the fractions are all recomputed so that no candidate has more than Q votes. If the number of elected candidates is less than k, the candidate with the fewest votes is eliminated (using a random process in the event of a tie) and its fraction is changed to 0. The votes of the eliminated candidate are thereby redistributed. The fractions and Q are then recomputed so that no candidate has more than Q votes, where Q is the sum of the votes of all remaining candidates divided by $(k + 1)$. Any candidate with Q votes is elected. The process of eliminating candidates, recomputation and declaring candidates elected continues until k candidates have been elected. The Meek rules are used by the Royal Statistical Society and have been endorsed by the ERS for computer-counted elections.

The rules proposed by Warren are the same as those proposed by Meek, except that instead of keeping the same fraction of whatever fractional value a vote has, an candidate keeps either the fraction p_i of a full vote or all of the remaining value of the vote, whichever is smaller.

The Problem of Sequential Eliminations

There is a further line of criticism to which all of the versions of STV discussed so far are subject. This line of criticism focuses on the fact that the versions of STV discussed so far all employ sequential eliminations. The problems that this causes can be seen most easily by noting that when any of the versions of STV discussed so far are used to select just one candidate, it reduces to the vote processing rule known as the alternative vote, or AV, discussed in Chapter 13. As shown in Chapter 13, AV lacks Condorcet consistency and non-negative responsiveness, which are highly valued properties that are possessed by a number of other RBVPRs for selecting one candidate. It might be expected that the most attractive RBVPR for selecting more than one candidate would reduce in a natural way to an attractive rule for selecting just one candidate.

David Hill has suggested that STV and AV might be modified so that a dominant candidate (calculated with respect to votes that had not yet been used to elect candidates) would be protected from elimination.[17] This would guarantee that when there was just one position to be filled, or when just one position remained to be filled, if there was a dominant candidate with respect to the remaining votes, that candidate would be selected. To me, however, this is too *ad hoc* to be satisfying. The manner in which the desired result is achieved for the case of one candidate to be selected leaves doubt as to whether the result for cases of selecting more than one candidate will be similarly attractive.

To avoid sequential eliminations, Michael Dummett suggested a procedure in which a search would be made for solid coalitions of a size that deserved representation, and when such a coalition was found, an candidate (or candidates) that the coalition supported would be selected.[18] If the solid coalition supported

more than one candidate, the candidate (or candidates) with the greatest "preference score" (Borda count) would be selected. Preference scores would also be used to determine which candidates would fill any positions not filled by candidates supported by solid coalitions.

I find Dummett's suggestion unsatisfying, for two reasons. First, suppose there are voters who would be members of a solid coalition except that they included an "extraneous" candidate, which is quickly eliminated, among their top choices. These voters' nearly solid support for the coalition counts for nothing, which seems to me inappropriate. Second, the Borda count seems to me an unattractive way of selecting candidates in general, because when it is used voters have a strong incentive to give artificially low rankings to serious contenders who they personally disfavor (as discussed in Chapter 13). This greatly compromises any meaningful interpretation of Borda counts.

Paired Comparisons by STV

In searching for a more attractive version of STV, I assumed that the answer would involve paired comparisons, so that it would reduce in a natural way to one of the rules in Chapter 13 that possess Condorcet consistency and non-negative responsiveness. However, the natural extension of paired comparisons, when more than one candidate is to be selected, is paired comparisons of *sets of candidates*, or "outcomes."[19]

Let I and J be two outcomes. Define X as the intersection of I and J. Define Y as the union of $I\backslash X$ and $J\backslash X$. (When I and X are sets, "$I\backslash X$" means "the elements of I that are not elements of X.") In words, X is the set of candidates that will necessarily be selected whether it is I or J that is chosen, and Y is the set of candidates whose selection depends on whether it is I or J that is chosen.

The number of votes for I over J can be conceived as the total number of voters whose first choice from Y is in $I\backslash X$, minus the number of those whose votes must be allocated to candidates in X to justify the fact that the candidates in X will be selected whether it is I or J that is chosen. This is a calculation very much in the spirit of STV. The beginning of the comparison would be just like STV. Each vote would be allocated to the candidate in the union of I and J ranked highest on that vote. Then a determination would be made of whether there were surpluses. However, the only surpluses that would be transferred would be those of candidates in X.

Voters in X would be pleased to have their votes transferred. As long as their votes were allocated to candidates in X, they would be exercising no real influence over the outcome: The candidates in X would be selected whether I or J was chosen. However, the procedure would not allow these voters to have any influence over the decision as to whether X would be supplemented by $I\backslash X$ or $J\backslash X$ until the candidates in X for which they voted had attained quotas.

Voters whose votes are allocated initially to candidates in $I\backslash X$ or $J\backslash X$ are in a different position. Such a voter does not want his or her vote transferred, because it could be that while the voter's first choice was in $I\backslash X$, his or her second was in $J\backslash X$, so that the transfer of part of the vote would go toward undoing the first choice. If the voter's second choice was in X, the transfer of a portion of the vote would still dilute its effectiveness. And if the second choice was also in $I\backslash X$, then the vote would contribute as much to the vote total for $I\backslash X$ when allocated to the first choice as when allocated to the second. Thus voters whose votes are allocated to candidates in $I\backslash X$ or $J\backslash X$ are pleased to have their votes not transferred.

After the transfer of all original and consequential surpluses of candidates in X, the vote total for I over J [the value of component (I, J) of the matrix of paired comparisons of outcomes], would be the sum of the votes for candidates in $I\backslash X$. A matrix of numbers generated in this way can be evaluated in any of the ways that a matrix of paired comparisons of individual candidates might be evaluated, including all the RBVPRs based on paired comparisons that are discussed in Chapter 13. Sequential eliminations are thereby avoided.

Because a vote might be transferred among several candidates in X before it was transferred to an candidate outside of X, it would be possible to use either the Meek rule or the Warren rule to decide how much of each vote to transfer. In the calculations that follow, the Warren rule is used. The process of comparing a pair of outcomes in the way described and identifying the vote total of each outcome will be described as a comparison of a pair of outcomes by STV, or CPO-STV. The matrix of such numbers will be referred to as a CPO-STV matrix of paired comparisons. The difference between such a matrix and its transpose will be referred to as a CPO-STV matrix of majorities.

Evaluating a CPO-STV Matrix of Majorities

Whatever rule a person believes is best for evaluating a matrix of paired comparisons or matrix of majorities when one candidate is to be selected, that rule can be applied to a CPO-STV matrix of majorities. The application of the chosen rule to the matrix yields an outcome just as in the case of the selection of one candidate.

To see how such a rule would be applied to the selection of more than one candidate, consider the set of rankings used previously to discuss alternative rules for transferring a surplus:

25	15	9	9	6	7	9
R	*R*	*S*	*S*	*T*	*U*	*V*
S	*S*	*R*	*U*	*U*	*T*	*T*
V	*T*	*T*	*T*	*R*	*S*	*U*
U	*U*	*U*	*R*	*S*	*R*	*R*
T	*V*	*V*	*V*	*V*	*V*	*S*

The first row of the CPO-STV matrix of majorities for this election is shown in Table 15.1.

Table 15.1 One row of a CPO-STV matrix of majorities

	{*R*,*S*,*T*}	{*R*,*S*,*U*}	{*R*,*S*,*V*}	{*R*,*T*,*U*}	{*R*,*T*,*V*}	{*R*,*U*,*V*}	{*S*,*T*,*U*}	{*S*,*T*,*V*}	{*S*,*U*,*V*}	{*T*,*U*,*V*}
{*R*,*S*,*T*}	—	1.86	9.65	31.0	29.0	28.0	33.0	31.0	30.0	42.0

Since this row consists entirely of positive numbers, outcome {*R*, *S*, *T*} is dominant and must be the winning outcome for any matrix evaluation rule that possesses Condorcet consistency.

As an example of how these numbers are computed, the comparison of {*R*, *S*, *T*} with {*R*, *S*, *U*} is made by excluding candidate *V* (the one candidate not in either outcome under consideration), transferring surpluses (if they exist, and they do) of candidates *R* and *S* (the intersection of {*R*, *S*, *T*} and {*R*, *S*, *U*}) and then taking the difference between the vote of *T* (that is, {*R*, *S*, *T*}\{*R*, *S*}) and *U* (that is, {*R*, *S*, *U*}\{*R*, *S*}). The vote at the final count for this comparison is

R 20.000;
S 20.000;
T 20.928;
U 19.072,

and the number in Table 15.1 for the comparison of {*R*, *S*, *T*} with {*R*, *S*, *U*} is the difference between 20.928 and 19.072.

The comparison of {*R*, *S*, *T*} with {*S*, *U*, *V*} is made by excluding no one, transferring the surplus (if it exists, and it does not) of candidate *S* (the intersection of {*R*, *S*, *T*} and {*S*, *U*, *V*}), and then taking the difference between the total votes of *R* and *T* (that is, {*R*, *S*, *T*}\{*S*}) and the total votes of *U* and *V* (that is, {*S*, *U*, *V*}\{*S*}). The vote at the final count for this comparison is:

R 40.000;
S 18.000;
T 6.000;
U 7.000;

V 9.000,

and the number in Table 15.1 for the comparison of $\{R, S, T\}$ with $\{S, U, V\}$ is the difference between 40 + 6 and 7 + 9.

CPO-STV reduces naturally, in the case of selecting one candidate, to an ordinary comparison of a pair of candidates. The sets of candidates to be compared are then the sets consisting of one candidate each. The comparison of candidate R with candidate S would consist of deleting from each ranking every candidate other than R and S, and then allocating each vote to whichever of R and S was ranked higher by that voter. CPO-STV is thus a natural extension of the idea of paired comparisons.

For cases in which the number of candidates is just one more that the number of positions to be filled, CPO-STV, when combined with any matrix-processing rule that is Condorcet consistent, yields the same result as ordinary STV. The reason for this is as follows: When there is only one candidate that will not be selected, the distribution of all surpluses leaves all candidates with exactly a quota of votes. The candidate that is not to be selected (assuming that there is no tie) will be the only one that does not achieve a surplus. Each comparison (by CPO-STV) of the set of selected candidates with an alternative set of the same size will involve picking one of the selected candidates, along with the defeated candidate, to not have their surpluses distributed. The selected candidate will thereby accumulate a positive surplus while the defeated candidate will accumulate less than a quota, so that each comparison will show the selected set defeating an alternative set, and the selected set will be a dominant outcome. Thus CPO-STV provides a natural bridge between methods of evaluating matrices of majorities for selecting one candidate, and STV for selecting one fewer than the number of candidates.

It seems reasonable to expect the processing of a CPO-STV matrix of majorities by a rule like the maximin rule to possess the property of $(k + 1)$-PSC, but I do not have a proof that it does.

Casual Vacancies and Structured Representation

A long-standing issue among advocates of STV is the question of the best way to fill a vacancy that occurs as a result of the departure of an elected candidate. One expedient that is sometimes used is to declare elected the unelected candidate with the highest vote total at the end of the initial count. However, it can be seen that this is not always appropriate. It could be that the voters who supported the candidate who departed were unanimous in regarding some other unelected candidate as their next choice after the departed candidate, in which case, clearly, that candidate ought to be awarded the vacant position.

It has seemed to me that, for hand counting or for the Meek and Warren rules, the most appropriate solution to the problem of casual vacancies is to first recount the election, taking note of just which votes are used to elect the departed candidate, and at what weight. One would also take note of the votes that, at the end of the count, had not been used to fill any quota, and their weights. (It would be possible for some of the votes in the first category to appear also in the second, in which case one would want to sum the weights from their two appearances.) The combination of these two categories of votes would constitute the votes that were not represented by elected candidates who had not departed, and that therefore deserved to be considered in the choice of a candidate to replace the departed one. Hence one would use whatever version of STV had been used in the initial election to select one candidate from these votes, taking account of their remaining weights.

With the advent of CPO-STV, however, a more elegant solution becomes available. One would simply form a matrix of paired comparisons among those outcomes that could be constructed by adding one unelected candidate to the previously elected candidates who had not departed, using all of the votes, and then evaluate that matrix by whatever matrix-evaluation rule was being used. This would identify the most representative outcome that was possible from the available candidates without requiring one or more of the sitting candidates to depart.

"Structured representation" is my name for the practice that is sometimes adopted, of requiring a representative body, elected by a procedure in the spirit of STV, to satisfy one or more criteria with respect to the distribution of representatives among categories. For example, an academic department might specify that its personnel committee shall consist of five tenured members and two non-tenured members. Previously, there has not been a satisfying way of remaining true to the spirit of STV while complying with structural requirements. With CPO-STV, however, there is a straightforward and attractive solution. Just confine the matrix of paired comparisons to those outcomes that satisfy the structural requirements, and evaluate it as any other matrix of paired comparisons would be evaluated. While this is conceptually simple, it is potentially explosive combinatorially. Therefore it remains to be determined whether it is practical.

Considerations of Computational Efficiency

The growing use of parallel processing and the growing speed of computers make computational efficiency not nearly as serious a problem for evaluating a CPO-STV matrix of majorities as it would have been just a few years ago. Nevertheless, because the number of possible outcomes can be exceedingly large (6,435 for selecting 7 of 15 candidates; over 184,000 for selecting 10 of 20 candidates), it will

be important to achieve computational efficiencies if the processing of CPO-STV matrices of majorities is to be workable.

One possible efficiency comes from guessing the answer, and then showing it to be such. If a row of a matrix of majorities has been computed and been observed to contain only positive numbers for the non-diagonal components, then the outcome corresponding to that row will be the winning outcome for any matrix-processing rule that possesses Condorcet consistency.

A second computational efficiency might be secured by beginning the computation with a "partial STV" count, stopping as soon as all surpluses have been distributed, and then giving a definite position to any candidate that had achieved a surplus. Then it would only be necessary to evaluate a matrix of outcomes that completed the set of selected candidates, which would often be a much smaller problem. One would hope that it could be proved that a practice of giving definite positions to candidates with initial or consequential surpluses would have no effect on the final result. It seems like it ought to be true, but I do not have a proof. Brian Wichmann (1988) has written a computer program that implements this rule for processing a CPO-STV matrix of majorities.[20]

A third computational efficiency, which seems likely to affect the outcome occasionally, though probably only in a very small percentage of cases, would be to search not for an outcome that dominated all others, but only for an initial outcome that dominated all outcomes that differed from the initial outcome just in the replacement of one of the initial outcome's candidates by an otherwise unselected candidate. I define an outcome with this characteristic to possess "local stability." If only local stability is sought in an outcome, the computational requirement is greatly reduced. I have written a computer program to implement this rule for processing a CPO-STV matrix of majorities. When outcomes are found to be in a cycle, this program focuses on the one whose worst loss is least bad. The search terminates if, after a check for local stability, the worst loss of the outcome that was checked is less bad than the worst observed loss of every outcome that the program considered in its search. The computational requirements of this program seem quite tolerable, so it will be important to undertake further research into the frequency with which this program yields answers that differ from those of Wichmann's more ambitious program.

For an election involving the selecting all but one candidate, the processing of a CPO-STV matrix of majorities by a rule that chooses a locally stable outcome is equivalent to ordinary STV, because local stability is global stability for the case of selecting all but one candidate.

For the selection of any number of candidates, consider a rule, R, that processes a CPO-STV matrix of majorities by starting with the result of the election by ordinary STV, and then changes one candidate at a time, searching for a locally stable outcome. R must possess $(k+1)$-PSC, by the following reasoning: The first estimate of the outcome, I, possesses $(k+1)$-PSC because the election is an STV election, and STV possesses $(k+1)$-PSC. If outcome J is chosen by R from the

union of outcome I and one other candidate, and if outcome I contains a set of candidates that must be included in the chosen set to satisfy $(k+1)$-PSC, then outcome J contains that set of candidates, since outcome J is chosen by STV. Since every estimate is chosen from the union of a previous estimate and one other candidate, by STV, every estimate contains any candidates that must be included to satisfy $(k+1)$-PSC. This proof that R possesses $(k+1)$-PSC could conceivably lead one to choose a rule like R over the more general processing of CPO-STV matrices of majorities, even apart from advantages of computational simplicity.

The Refinement–Comprehensibility Trade-off

Each refinement in STV answers a distinct objection to an earlier version, but at some cost in making the rule less comprehensible. So it is a serious question as to whether all refinements are worth their cost in comprehensibility. Some members of the Electoral Reform Society who are concerned with spreading the acceptance of STV (and have considerable experience with trying to explain it) believe that any rules more complex than those introduced in the early seventies for elections in Northern Ireland (a variation on the Senatorial Rules) would be too incomprehensible to be acceptable to a general electorate. These people reject such refinements as Newland and Britton's introduction of a rule that declares elected an candidate with fewer votes than a quota, but enough to be guaranteed election in view of the number non-transferable votes that have been encountered. Others would stop at the NB rules, or the Meek rule. However, whatever the tolerance of complexity may be, the more complex rules provide valuable insight into the cost of overcoming limitations of the simpler rules, and experiments with them can reveal how often the more complex rules yield different outcomes, and therefore the losses, if any, from using simpler rules.

A Recommendation for a Hybrid System

Having considered vote processing rules for selecting more than one candidate from a variety of perspectives, it is now appropriate for me to make a recommendation, taking account of the technical properties of the various rules, as well as the difficulty of understanding rules, the cost of voting and the cost of computation.

Should a rule for selecting more than one candidate be a system of proportional representation? Or should single-member constituencies be used, to reduce fragmentation in an elected body? It is true that single-member constituencies greatly reduce the representation of minorities, but they also tend to generate geographical polarization in elected bodies. I believe that geography should not be singled out as the one characteristic by which people are identified for purposes of

representation. I also believe that elected bodies can and should be organized in ways that respect the rights of minorities without permitting minorities to block action, and that the potential problem of shifting majority coalitions should be handled by making decisions that are binding for suitable periods of time. Thus I favor systems of proportional representation.

Should a system of proportional representation be of the sort that Tullock proposed, in which each voter can designate his or her representative? If the only purpose of an elected body were to represent the interests of voters, this might be a good idea. However, part of the purpose of an elected body is to exchange views and debate, creating greater consensus about what courses of collective action are best. For this purpose it is good to have several representatives rather than a single representative with greater weight. Therefore I endorse the idea of a system of proportional representation with a fixed number of representatives, each with equal voting weight.

The next question is, should a system of proportional representation be of the party-list type or the STV type. One virtue of systems of the party-list type is that it is very easy for voters to vote under such systems, and reasonably easy to count the votes. In my judgment, however, these virtues come at too high a cost. Under party list systems voters are able to convey very little about their preferences. For this reason, I regard voting systems of the STV type to be superior to systems of the party-list type. Nevertheless, it is important to note that, under STV, there is typically only a small minority of voters who rank all of the available candidates.

For this reasons, I favor a hybrid of STV and party-list type systems. Voters should be permitted to rank a maximum of perhaps ten or twelve candidates, and then to name a party whose list would receive any residual value of the voter's vote. It would be much easier to implement computerized voting if there were a limit on the number of candidates that a voter could name. Few voters would be constrained by the rule, and even fewer of them would be noticeably disadvantaged, since there is typically only a very small fraction of the initial weight of a vote left after it has been passed through ten or twelve candidates.

I would recommend that votes be counted by the Warren variation on the Meek rules for computer counting. The Meek rules for computer counting are much faster and more elegant than the hand-counting rules. I favor the Warren variation because this rule accords with my conception of fairness in allocating the "cost" of electing a candidate among the candidate's supporters. I would not recommend the CPO-STV rules yet because they are still too experimental, although they may eventually prove to be attractive.

The next chapter considers the problem of assigning endogenous weights to voters, according to the intensities of their preferences.

Notes

1 Gordon Tullock, *Toward a Mathematics of Politics* (Ann Arbor, 1967), pp. 145–6.
2 For a detailed discussion of party-list systems of proportional representation, see Enid Lakeman, *How Democracies Vote: A Study of Electoral Systems* (London, 1974).
3 These are very ably reviewed in Michael Balinski and H. Peyton Young, *Fair Representation: Meeting the Ideal of One Man, One Vote* (New Haven, 1982).
4 Ibid., pp. 84–93.
5 Ibid., pp. 23–35.
6 Ibid., p. 35.
7 Ibid., pp. 18–22.
8 Ibid., p. 157.
9 This fact and others in the discussion of STV that follows are based on Clarence Hoag and George Hallett, *Proportional Representation*, (New York, 1926); Robert Newland and Frank Britton, *How to Conduct an Election by the Single Transferable Vote*, (London, 1976); and I. D. Hill, "Some Aspects of Elections—to Fill One Seat or Many," *Journal of the Royal Statistical Society, Series A* (General) 151/2 (1988): 243–75.
10 Dummett, *Voting Procedures*, p. 282.
11 Newland and Britton, *How to Conduct an Election by the Single Transferable Vote*.
12 Dummett, *Voting Procedures*, p. 283.
13 Brian Meek, "Une nouvelle approache du scrutin transférable", *Mathématiques et Sciences Humaines* 25 (1969): 13–23.
14 I.D. Hill, B.A. Wichmann and D.R. Woodall, "The Single Transferable Vote by Meek's Method," *Computer Journal* 30 (1987): 277–81.
15 C.H.E.Warren, "Counting in STV Elections," Unpublished paper, 19 Rectory Road, Farnborough, Hants GU14 7BU, England (August 1983).
16 There is a fourth, more technical feature. When some of the votes due to be transferred cannot be transferred, the number of votes required for a candidate to be deemed elected is reduced by the ratio of the number of non-transferable votes to one plus the number of positions to be filled minus the number filled by candidates whose surpluses have already been transferred. That is, the reduction in usable votes is distributed over one more than the number of candidates whose final scores are still variable.
17 Hill, "Some Aspects of Elections," p. 260.
18 Dummett, *Voting Procedures*, p. 284.
19 See Nicolaus Tideman, "The Single Transferable Vote," *Journal of Economic Perspectives* 9/1 (1995), pp. 27–38, esp. pp. 36–7 and Nicolaus Tideman and Daniel Richardson, "Better Voting Methods through Technology: The Refinement-Manageability Trade-off in the Single Transferable Vote", *Public Choice* 103 (2000) pp. 13–34, esp. pp. 28–33.
20 Brian Wichmann, untitled computer program written in 1988, available from Wichmann at 5 Ellis Farm Close, Woking, GU22 9NQ, England. Ask for the program that counts votes by Tideman's method.

This page intentionally left blank

Chapter 16

Vote Processing Rules With Endogenous Weights for Self-Interested Advocates: Revealing Intensities of Preferences

Introduction

When there are just two options under consideration, so that cycles are not an issue and majority rule is well defined, there is a tendency to think of the problem of voting as one of simply determining what the majority wants. But it is not necessarily true that the majority should always prevail. It is possible that the aggregate intensity of the feelings about a decision on the part of the minority is greater than the aggregate intensity of the feelings on the part of the majority, so that it is arguable that the minority ought to prevail. The rationale for majority rule from the principle of insufficient reason that was developed in Chapter 8 relied on the presumption that there was no basis for asserting that the average intensity of feeling among the minority was any greater than the average intensity of feeling among the majority. Thus if there is some basis in a particular election for asserting that the aggregate intensity of the feelings of the minority is greater than the aggregate intensity of feelings of the majority, then it can be argued that in that particular instance the minority should prevail.

But what is meant, operationally, by "the aggregate intensity of feeling?" Should voters be asked, on a scale of 0 to 100, how strongly they feel about the matter at hand? It is doubtful that useful information would emerge from such a practice. Even if people were moral enough to resist the temptation to report an intensity of 100 every time, there would be a problem of ascertaining what was meant by intensity. How would people know the true intensities of their own preferences? The meaning that economists generally give to "the intensity of a preference" is the value to a person, in money, of having one outcome instead of another. We will start with that meaning and consider some alternatives later.

Imagine asking voters, "How much would you be willing to pay to have one outcome instead of another?" That at least is a question for which it is reasonable to suppose that each voter knows the true answer. But is it reasonable to suppose that people will give those true answers if asked?

That depends in part on what action is to be taken on the basis of the answers. If willingness to pay money is regarded as the appropriate measure of intensity of preferences, and if people answer the question honestly, then the collective

decision rule that is implied is that whichever option is more highly valued should be chosen.

But would those who prevail be required to pay in accordance with the amounts they named? If they were not required to pay, then it would be reasonable to be concerned about the possibility that people would show no restraint in the amounts they named, and the whole exercise would be meaningless. But if those who prevail were required to pay according to their expressed preferences, then voters might calculate that whatever was going to happen was unlikely to be influenced by whatever amounts they named individually, so they might as well say that they were not willing to pay anything to influence the decision, and save their money. A few people might name positive amounts out of a sense of duty, out of miscalculation, or out of speculation that few others would say they were willing to pay anything, but it would be unreasonable to suppose that the overall outcome would have any systematic relationship to actual intensities of preferences.

Despite these considerations, there is a tradition, mostly outside of economics, of taking seriously the responses that people offer to surveys about the value of public goods. This tradition gives the name "contingent valuation" to values for public goods developed in this way, after the idea that people are reporting their valuations of contingencies with respect to public goods. Economists generally regard such valuations as unreliable.[1] They tend to respect only systems that work well when people are regarded as responding to their personal incentives.

There are two types of devices that economists have suggested for motivating people to report the intensities of their preferences honestly: ones that rely on uncertainty in the outcome and ones that employ marginal cost pricing. These are compared in the sections that follow.

Choosing between Two Options

Consider the simplest situation in which the problem of taking account of intensities of preferences arises, a collective decision between two options that have no resource costs. To make the example concrete, suppose the question is whether a town should adopt a plan of one-way streets. Suppose it is agreed that the people whose views are relevant are the citizens of the town, and that the plan ought to be adopted if the aggregate intensity of the preferences of those who favor the plan exceeds the aggregate intensity of the preferences of those who oppose it. Then the problem is to motivate the citizens to report the intensities of their preferences honestly.

A Mechanism that Relies on Uncertainty in the Outcome

Earl Thompson proposed a mechanism that relies on uncertainty in the outcome.[2] While he called his mechanism the D-process, I will follow Martin Bailey[3] and call

it the Thompson insurance mechanism. It can be described as a government-run system of insurance against unfavorable outcomes, with the insurance market used to make collective decisions as well.

Suppose that everyone agrees that there is a 75% chance that the proposed system of one-way streets will be approved. Then, for a person who places a value of $60 on implementation of the system, $15 is an actuarially fair price for insurance against the loss that he would suffer if the system was not implemented. Similarly, for a person who would experience a lost of $60 from implementation of the system, $45 is an actuarially fair price for insurance against the loss that he would suffer if the system was implemented. If everyone bought full insurance against their potential losses, then it would be possible to discover whether the system of one-way streets was worthwhile overall, by comparing the losses from implementing the system with the losses of not implementing it. If the option with the smaller losses is chosen, then the collected insurance premiums will be more than enough to pay the losses. For example, if those who favor the one-way streets buy insurance against $5,000 in losses from maintaining the status quo (paying $1,250) and those who oppose one-way streets buy insurance against losses of $6,000 from implementing one-way streets (paying $4,500) then the sum of all premiums ($5,750) is more than enough to pay the losses (of $5,000) of the less-valued option. In general, if losses of L_A are claimed from adoption and losses of L_R are clamed from rejection, then, for any probability of adoption, p, the sum of the premiums, $pL_A + (1–p)L_R$, will be greater than or equal to the minimum of L_A and L_R, since it is a weighted average of the two. Thus a rule of selling actuarially fair insurance for all potential losses and implementing the option with the smaller losses will yield a choice for the more efficient option and also yield a budget surplus (or a balanced budget in the case of a tie.)

Whether the Thompson insurance mechanism would motivate people to report the intensities of their preferences honestly depends on their beliefs and their preferences. A person who believes that the announced probability of adoption is true and places no positive value on taking risks will be motivated to report his valuation honestly. Whether a person who believes that the probability of adoption is different from what is announced will be motivated to report his valuation honestly depends on the relationships among the announced probability, p^*, of the outcome that the voter believes is underpriced, the voter's belief about that probability, p, and the weight, w, that the voter places on dollar lost compared to a dollar gained. If w is greater than or equal to p/p^*, then the voter is motivated to report truthfully. Otherwise, some misstatement of his preference will yield a greater subjective utility than the truth.

Thompson assumes that the government would be the only source of information about the probability that any proposal would be approved. To ensure that people would believe the probabilities announced by the government, he proposes that the government use "sampling devices, polls, propaganda, or appropriate taxes on the behavior of those bold individuals who indicate some

consistent pattern of disagreement with the government probability estimates."[4] Martin Bailey proposed that the probabilities for the Thompson insurance mechanism be determined by the application of statistical sampling theory to the responses of a large, randomly chosen representative legislature that would develop the proposals.[5] Bailey proposed that in evaluating proposal, the legislature would use the mechanism described in the next section.

A Mechanism Based on Marginal Cost Pricing

Edward Clarke proposed a mechanism based on marginal cost pricing.[6] The essence of Clarke's idea is that people will be motivated to report the intensities of their preferences for public goods honestly if the price that each person is charged for participating in a collective decision is the marginal cost of taking account of that person's preference.

What is the marginal cost of taking account of a person's preference? If the decision would be made the same way with or without the person's participation, then the marginal cost is zero. But if the decision would be made differently if the person had not participated, then there is generally a positive marginal cost of taking account of that person's preference.

Continue with the example of one-way streets. Suppose that each voter in a town is asked how much he or she would pay to have a plan of one-way streets implemented, or to ensure that it would not be implemented. Suppose that the total amount stated as the value of one-way streets to those who want them is $5,030, while the total amount stated by those who oppose the plan is $5,000. And suppose that one person stated that it would be worth $50 to him to have the plan introduced. Then this person is "pivotal." With his participation the plan is seen to be worth $30 more to its proponents than the cost to its opponents, but without his participation it would be seen to be worth $20 less to its proponents than its cost to its opponents. The marginal cost of attending to the preference of the person who places a value of $50 on the system of one-way streets is the $5,000 value lost by the opponents if the system is introduced, minus the $4,980 value received by the other proponents of the system, for a net marginal cost of $20. More generally, if a person is pivotal (that is, if the person is from the side that prevails in the collective decision and states a value that is greater than the winning margin), then the marginal cost of attending to that person's preference is what the winning margin would be, in the other direction, without that person's participation. An alternative formulation that produces the same answer is that the marginal cost of attending to the preferences of a pivotal person is the amount of money offered by that person, minus the winning margin.

Gordon Tullock and I gave Clarke's idea the name, the "demand-revealing process."[7] It is also known sometimes as the "pivotal mechanism." Tullock and I also named the amount that a pivotal voter is required to pay a "Clarke tax."

When the demand-revealing process is used, no individual can improve his or her position by misstating the intensity of a preference, and, depending on the

preferences of others, he or she may be worse off. To see this, consider the person who stated that he placed a value of $50 on the system of one-way streets, and suppose that $50 is his true valuation. If he states a value of more than $50 he is neither better off nor worse off. The winning margin in the other direction without his participation remains $20, and that is his Clarke tax. If he names an amount between $20 and $50, then the system of one-way streets continues to be approved, and he continues to pay a Clarke tax of $20. Now suppose that he names an amount less than $20, say $15. In these circumstances there is more measured opposition to the plan than support for it, and it is not adopted. The person who truly feels that the adoption of the plan would be worth $50 to him has passed up the opportunity to obtain that benefit for a cost of only $20. He has suffered a $30 loss, compared to the consequences of stating his preferences honestly. The same is true if he states that he opposes the plan when he actually favors it.

Now consider the possibilities for a person on the losing side. Suppose that among those who opposed the system of one-way streets there was a person who truthfully reported that he would be willing to pay $40 to block the introduction of the system. Because of the magnitude of the support for the system, he and others who feel as he does are thwarted. If this person names an amount less than $40, or falsely states that he favors the system, he remains thwarted. The system is approved by an even larger margin. If he names an amount between $40 and $70 the system continues to be approved, but by a smaller margin. If he names an amount greater than $70, then the plan will be defeated, but he will be charged a Clarke tax equal to what the margin in favor would have been without his vote, namely $5,030 – ($5,000 – $40) or $70. Thus it has cost this person $70 to achieve a result that is worth $40, and he is $30 worse off than he would have been if he had responded truthfully. These arguments do not depend on the particular numbers used. Under the demand-revealing process there is no possibility for an individual to make himself or herself better off by misstating the intensity of his or her preference.

While no *individual* can make himself better off by a misstatement of the intensity of an individual preference, there is a chance for a *coalition* to make themselves better off by coordinated misstatements of the intensities of their preferences.

Suppose that there are two opponents of the plan of one-way streets who are each willing to pay $40 to defeat the plan. If they were to agree that they would both state $80 as the intensity of their preference, then there would be an additional $80 in the stated value opposed to the plan, and it would be defeated by $50. Each member of the coalition would have to pay a Clarke tax of $30 (the winning margin in the other direction without one of the strategic votes), but this would be less than the $40 true value to each of them of defeating the plan. If they were to increase their stated values to more than $110, then the winning margin in their favor would be so large that the plan would continue to be defeated when either of their votes was removed, so they would pay no Clarke taxes.

The significance of the possibility for coalitions to gain by misstating their preferences is diminished by the difficulty of guessing what misstatement is necessary, and the possible losses from guessing wrong. If the two opponents both named $110 as the value each would place on blocking the plan, making the total amount in opposition $5,140, and if the amount voted in favor of the plan were $5,130 rather than $5,030, then each member of the coalition would wind up paying $100 for something that was worth only $40 to him.

The opportunity for coalitions to gain is further diminished by the fact that every member of the coalition has an incentive to state his or her preference honestly rather than abide by what the coalition agreed. No matter what the other members of the coalition do, each member of the coalition can make him or herself worse off but not better off by misstating the intensity of his or her preference. This incentive to disregard any coalition agreement is particularly telling when the amount that each person votes can be kept secret. Such secrecy could be achieved by requiring voters to have accounts from which their Clarke taxes would be deducted. If no paper records of the amounts in such accounts were ever provided to voters, then no one would ever be able to prove to his or her coalition partners that a vote had been cast in the agreed fashion, so the coalition partners would also be unable to prove that someone had not voted in the agreed fashion. This should make it reasonably difficult to distort outcomes through coalition agreements.

Even without a coalition agreement, voters may value a project because of the benefits they believe others would receive from it. Howard Margolis has suggested that people would treat opportunities to vote under the demand-revealing process as opportunities to express civic-mindedness by offering sacrifices to see that the right thing was done, rather than opportunities to report the benefits that they would receive from having decisions made in the ways that they preferred.[8]

This raises the question of whether reports of such "secondary value" would constitute inaccuracies in the aggregate value of a project. The answer depends on what is meant by "value." Traditionally, economists have meant by "value" the amount of one thing that a person is willing to sacrifice to acquire another thing. The reason for the willingness to sacrifice is not questioned. From this perspective, secondary value is as valid a component of aggregate value as primary value. Its inclusion does not constitute a distortion in the measurement of aggregate value.

Turn now to the question of what is to be done with the Clarke taxes. If they are shared equally among the voters, then a voter who expects to be on the losing side acquires an incentive to overstate the intensity of his preference, so that the outcome would be closer and more Clarke taxes would be collected from voters on the winning side, increasing the amount that he would receive. A voter who expected to be on the winning side would have an incentive to understate his preference for the same reason. To avoid these incentives, the Clarke taxes can be wasted (used to purchase something produced outside the town, which is then discarded), or two towns can agree that each will give the other any Clarke taxes

collected, or a town can hold an auction among non-citizens, prior to any election, for the right to receive the Clarke taxes. However, if a collectivity has even a few hundred voters, it is unlikely that there would be any perceptible misincentives from simply dividing Clarke taxes equally among the voters.

Bailey has pointed out that the expected value of the budget imbalance in the demand-revealing process can be reduced by an order of magnitude while maintaining incentives for truthful reports, by assigning to each voter, i, a lump-sum transfer equal to $S_{-i}(N-1)/(N-2)^2$, where S_{-i} is what the surplus would have been if i had abstained, and N is the number of voters.[9]

The Possibility of a Tie

Both for the Thompson insurance mechanism and for the demand-revealing process, there is a possibility of a tie. If there is a tie, then the available evidence provides no basis for choosing between the two options. One would reasonably flip a coin or use some other random process to choose between the options. For the Thompson insurance mechanism, the only difference would be that there would be no surplus. The collected premiums would be exactly enough to pay the losses of those whose disfavored option was chosen. For the demand-revealing process, a tie would mean that every voter was pivotal and would be required to pay *all* of the sacrifice that he or she offered for the chosen outcome.

The Units in which Value is Measured

Both for the Thompson insurance mechanism and for the demand-revealing process, dollars are somewhat unattractive as a measure of value, because there is no reliable measure of the marginal utility of dollars to different individuals, and there is substantial consensus that dollars do not have the same marginal utility to all persons. This is a concern because, ideally, one would want to measure intensities of preferences in terms of utility.

The demand-revealing process does not require that value be measured in dollars. Any measure that represents a sacrifice for individuals and can be added across persons could be used.[10] One could measure sacrifice by a quantity of effort, such as the number of multiples of a person's body weight that he or she was willing to carry (say in bricks) the length of a football field. But this would discriminate against the frail and the overweight. One could make the measure of sacrifice the number of hours that one was willing to spend picking up roadside trash, but, again, it is unlikely that all people find an hour of this activity equally costly. One could weight the dollars offered by different people according to estimates of the marginal utility of dollars to them, but there is likely to be substantial disagreement about how these estimates should be constructed. If the dollars were to be weighted, one could think of each vote as representing a certain number of "utils," which would be calculated by multiplying the dollar vote by the estimate of utils per dollar for the voter. The outcome would be determined by

counting the total utils on each side. Clarke taxes would be calculated initially in utils, and then converted to dollars for payment.

The problem of the most attractive measure of sacrifice remains a difficult issue for the demand-revealing process. Dollars (possibly weighted) and hours of community service are the measures of sacrifice for the demand-revealing process that seem most likely to be most attractive.

Since the Thompson insurance mechanism collects the value reported by voters and uses it to pay insurance claims, it would not be feasible to operate this mechanism on the basis of a measure of sacrifice, such as carrying bricks, that did not generate revenue with which to pay insurance claims. If one attempted to operate the Thompson insurance mechanism with dollars weighted by the marginal utility of dollars, there would be a possibility of a revenue shortfall whenever the weighted average of the marginal utility of money on the winning side (weighted by amount of insurance bought) was less than the weighted average of the marginal utility of money on the losing side. Still, it would be reasonable to expect to be able to use the Thompson insurance mechanism with dollars weighted by the estimates of the marginal utility of money if the mechanism was going to be used for many decisions and there was no reason to expect that approved proposals would generally favor those with low marginal utilities of money. Surpluses from some applications would offset deficits from others, and almost all applications would generate some surplus when measured in utility. The Thompson insurance mechanism could also be used if there was some source of revenue from which any such deficits that arose could be financed.

The Possibility of Compensation

Even with the best possible measure of sacrifice, a collective decision that is not unanimous represents a loss for some persons. If a proposal has benefits that exceed its costs, then there is some variation on the proposal that, by adding charges on those who benefit and compensation for those who lose from the proposal, makes everyone better off. If the situation prior to the decision about the proposal is just, then the justice of requiring those who lose to accept their losses is questionable. The reason for requiring those who lose to accept their losses, when the situation prior to the decision is just, is that we want to live in collectivities that make some binding collective decisions without the cost of achieving unanimity, and this entails imposing uncompensated losses on some persons.

It is ironic that the mechanisms discussed in this chapter, when they work as intended, reveal exactly what compensation would have been necessary to turn proposals with positive net benefits into ones that would have provided positive benefits for everyone, and yet if there were any suspicion that the revealed information would be used to determine compensation, then people would lose the incentive to report their preferences honestly. It is not possible to motivate voters to provide the information that is necessary to determine appropriate compensation if voters suspect that the information may be used to determine compensation.

Consider the Thompson insurance mechanism. If you are asked how much insurance you want to buy at a price of 10%, against adoption of a proposal that will truly cost you $1,000, and if you believe that raising your reported loss by $1 will raise the compensation you receive by x, then the cost of insurance is not 10%, but rather 10% minus x. With insurance underpriced, you have an incentive to buy too much. Similarly, if under the demand-revealing process, reporting a willingness to pay $1,000 to block a proposal results in compensation of some positive amount x, then it pays to report a loss of $1,000 when your actual loss is $1,000 minus x. The incentive to report intensities of preferences honestly is lost.

One might think that, while incentives are distorted if compensation is linked to reported preferences, observed correlations in preferences could be used to determine compensation after the vote had occurred, if each voter's response was removed from the information used to determine his or her compensation. Unfortunately, this distorts incentives too. If, under the demand-revealing process, a voter is asked to evaluate a proposal that will cost him $1,000, and he knows that if it is approved it will be followed by compensation which, in his subjective opinion, has an expected value of x, then the expected cost of the proposal to him is not $1,000 but rather $1,000 minus x. The capacity of the process to determine the pre-compensation cost is lost. The Thompson insurance mechanism is affected in the same way by ex post compensation.

The only way that compensation can be incorporated into voting rules with endogenous weights, such as the Thompson insurance mechanism and the demand-revealing process, is to include the best available estimates of the compensation for losses in the original proposal. If, prior to the vote on the proposal, there is consensus that persons in some category will lose from the proposal, then the proposal should be amended to include compensation for them. If there is consensus that one category of voters will benefit disproportionately, then it is reasonable to require more from them.

One possible way of assigning compensation would be to appoint an assessor (outside the group making the decision) with the power to set compensation, and with incentive pay that depended significantly on achieving a small dispersion in average reported benefits across groups for which it was legally permissible to vary compensation. The ideal against which the compensation assessor's work would be judged would be that in every legally distinguishable group, the average per capita reported net benefit of the proposal would be the same. It would generally not be legally possible to set compensation perfectly, because, in order to keep compensation from being arbitrary, we would require it to be a function of specific facts about voters (for example, income, wealth, residential and business location, family composition and age).

It would be reasonable to expect that, if a society made a concerted effort to incorporate into proposals the compensation that would made the proposals as close to Pareto improvements as possible, then the dispersion in net benefits would

diminish over time, as information accumulated about the correlates of reported benefits and costs.

Bailey discusses the possibility of weighting harms more heavily than benefits in deciding whether to adopt a proposal.[11] This bias in favor of the status quo would be consistent with utility maximization if the loss of utility that a person typically experiences from the adoption of a measure, when he or she would pay $100 to stop it, is noticeably greater than the increase in utility that the same person experiences from the adoption of a measure, when he or she would pay $100 to ensure its adoption. There is evidence that this is true.[12]

Intentionally Redistributive Measures

The evaluation of intentionally redistributive measures by either the Thompson insurance mechanism or the demand-revealing process raises new issues. When a proposal is intentionally redistributive, compensation is not relevant in the way that it is for other proposals. For example, additional transportation vans might be proposed for handicapped persons, with the intention that handicapped persons would benefit from the adoption of the proposal. Thus handicapped persons would not be expected to pay the costs of providing the vans, making the proposal intentionally redistributive.

Among proposals that are intentionally redistributive, it is useful to distinguish between those that are intended to be approximate Pareto improvements and those whose justification is that the benefits to the beneficiaries outweigh the costs to the payers.

A proposal is an approximate Pareto improvement if the proposal is valuable enough, and the payers care enough about the recipients, that the payers as a group receive positive net benefits when the consequences for different payers are weighted the same way as for other proposals. Then one can say that, while there is some unintentional redistribution arising from the imperfect assignment of costs, the proposal has the potential for benefiting all persons, apart from such errors. If the errors are expected to be random, one can imagine getting unanimous agreement, at a constitutional stage, that such measures should be approved when the net value to payers is positive. Either the Thompson insurance mechanism or the demand-revealing process can be used to evaluate proposals in this category, provided that the votes of the recipients of redistributive benefits are separated from the votes of payers, and approval by both groups is required. If the two types of votes were combined, measures would pass when the benefits to recipients more than offsetting the costs to payers. But in this case one could not say that the proposal would have benefited everyone apart from random errors in the assignment of costs, because the recipients are not eligible to have their cost shares increased. Thus to ensure that the only measures that pass are ones that would benefit everyone, apart from unintentional errors in the assignment of cost shares, the votes of recipients must be separated from the votes of payers, and a proposal must receive the approval of both groups to pass.

Now consider redistributive measures that are not proposed as approximate Pareto improvements, but rather are intended to cause losses to payers and are defended on the ground that the gains to recipients will more than offset the costs to the payers. If given a choice, the typical payer would not permit such measures to be considered. Other things being equal, payers would choose to live in jurisdictions that did not allow such measures. Thus such measures do not satisfy the definition of a good collective decision developed in Chapter 6. If someone who has the power to impose use of a vote-processing rule with endogenous weights wishes to impose it nevertheless, for the evaluation of measures that are intended to cause losses to payers, the following analytic points are relevant.

First, unlike the case of measures that are intended as approximate Pareto improvements, in this case there is no rationale for separating the votes of payers and recipients.

Second, there is no prospect of equating cost shares with benefit shares. In other applications of decision rules with endogenous weights, a systematic discrepancy between cost shares and benefit shares can be treated as a failure in the process of assigning cost shares. However, no such inference is appropriate when the proposal is intended to be redistributive.

Third, the Thompson insurance mechanism becomes problematic in a new way. If there is a 30% probability that a proposal will be approved, then, under the Thompson insurance mechanism, beneficiaries would need to face insurance premiums equal to 70% of the benefits that they would receive, to be motivated to report accurately the benefits of the proposal to them. And since beneficiaries would be assigned zero cost shares, the benefits to them, and therefore the costs of insurance would be substantial. While it would still be *possible* to use the Thompson insurance mechanism, this undoing of the intended redistributive effect of proposals would seriously reduce its effectiveness. This problem could be ameliorated by considering only proposals that had a high probability of being approved.

A related problem arises with the use of the demand-revealing process to evaluate intentionally redistributive proposals that are intended to harm payers. Where the problem with the Thompson insurance mechanism arose to the extent that proposals were unlikely to pass, the problem with the demand-revealing process arises when proposal pass with margins that are so small that there are substantial Clarke taxes. Then, once again, those who are intended to benefit from proposals can wind up paying all of their benefits in Clarke taxes. This problem can be ameliorated by specifying that any Clarke taxes collected from a recipient of benefits will be refunded to other recipients.

If there is a justification for projects that intentionally harm some persons for the benefit of others, it is that justice requires the projects. The intensities of preferences are not relevant to the evaluation of such projects. I recommend that neither the Thompson insurance mechanism nor the demand-revealing process be used for the evaluation of such projects.

Are Incentives Identified Accurately?

The claim that voters have incentives to report the intensities of their preferences honestly, under vote-processing rules with endogenous weights, rests on the assumption that what concerns voters is the benefits and costs that will flow from the decision to be made, and not any satisfaction from the act of voting, or from other consequences of voting. In an experiment in which fraternities made decisions by the demand-revealing process, I found that, when interviewed after a vote, voters sometimes said that they had reported intensities of preferences that exceeded their actual intensities, for the thrill of taking a risk and seeing if they could get away with it.[13] Under the Thompson insurance mechanism, people who enjoyed risk might try to guess which way the vote would go and then claim a loss from that outcome, just for the thrill of betting.

There are at least two other motivations, besides an enjoyment of risk, that might lead people to report intensities of preferences that differed from the satisfaction they would receive from a potential outcome. First, they might get satisfaction from knowing that they had offered money in support of the right choice. (For the Thompson insurance mechanism, they lose the money if their stated conception of the right choice prevails, but it comes back with more added on if their conception does not prevail; for the demand-revealing process, they lose only such a part of the money as makes them pivotal.) Second, under the demand-revealing process, they might experience satisfaction or loss from changes in the expected income distribution that would be caused by the changes in expected Clarke taxes paid by others as a result of their votes.

A Combined Mechanism

Bailey suggested that proposals for governmental action should be evaluated by a combination of the demand-revealing process and the Thompson insurance mechanism.[14] He argued that the incentive to vote is stronger under the Thompson insurance mechanism, since the purchase of insurance has consequences for voters in all cases, while the reporting of a preference under the demand-revealing process has consequences only when voters are pivotal. On the other hand, he notes that Calfee and Rubin explained that voters have no incentive to buy insurance against disfavored outcomes that offend them morally rather than harming them financially, because insurance improves utility only when the event insured against increases the marginal utility of money.[15] Therefore Bailey proposes that voters be allowed to both buy insurance through the Thompson insurance mechanism and also report prospective losses that they have not insured against. The outcome would be determined by the total value reported through both mechanisms. Insurance claims would be paid according to insurance bought through the Thompson insurance mechanism, and Clarke taxes would be paid by voters who reported uninsured losses greater than the winning margin.

While each mechanism operating by itself will never have a deficit, it is possible for the two operating together to have a deficit. This occurs when more losses than gains for the winning option are reported through the Thompson insurance mechanism, but these losses are more than offset by the fact that more gains than losses for the winning option are reported through the demand-revealing process. Then the insurance mechanism operates at a loss, and there will generally not be enough Clarke taxes to offset this loss. It is still possible to use the combination mechanism, as long as there is some other source of financing from which to pay any deficits that arise. The problem of possible deficits is smaller when the combined mechanism is to be used for many decisions, so that surpluses from some decisions can offset losses from others.

The Problem of Future Uncertainty

A problem that arises both with the Thompson insurance mechanism and with the demand-revealing process is that people may have different ideas about later consequences of a choice. Suppose that a proposal is made that a system of one-way street be tried for a year, with a decision on continuation of the system at the end of the year. Under the Thompson insurance mechanism, voters would evaluate the proposal in terms of its costs or benefits to them for one year, plus the expected net benefit of the decision to be made after a year. To attain commensurate measures of the value of the proposal, all voters would need to use the same probabilities about what would happen after a year. It is hard to know how this could be achieved. Any dimensions in which people have disparate beliefs about future implications of a proposal leads to non-commensurate measures of value, and therefore to possible inefficiency.

Under the demand-revealing process, there is no insurance market to mediate beliefs about the action to be taken after a year, but it remains true that achieving commensurate measures of value requires that all voters have the same belief about the probability that the system of one-way streets will be continued after the first year. If some evaluate it under the assumption that it will be terminated a year while others evaluate it under the assumption that it will last indefinitely, then the aggregate reported value will have no economic meaning. Any other difference in beliefs about the future implications of a decision will have a similar effect.

Incorporating Resource Costs

The example of one-way streets was used for the initial illustration of voting rules with endogenous weights because the resource costs of implementing one-way streets are small compared to the value of other consequences, and an example without resource costs makes it easier to focus on the central characteristics of the systems. Now suppose there is an option under consideration with significant resource costs, say a plan to install traffic lights in a town. Then the question of how the costs will be paid must be addressed along with the question of whether the project is worthwhile. One cannot use a voter's report under the Thompson

insurance mechanism or the demand-revealing process to assign a cost to him or her, because that would destroy the incentive to report intensities of preferences honestly. To maintain incentives for honest reporting of preferences, the costs of any proposal must be allocated among voters in a way that is prescribed before the decision is made. Costs can be allocated through any traditional tax. Alternatively, as with compensation for an option without resource costs, a collectivity can use an assessor who is rewarded for assigning costs in such a way that all groups that are distinguishable for tax purposes have the same average net benefits. Once an allocation of costs has been determined, a collectivity decides whether to implement an option with specified financing in the same way as with a proposal that has no resource costs.

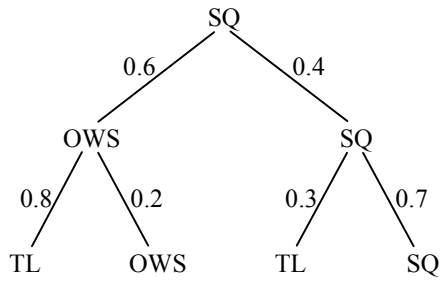
Choosing among Three or More Options

Now suppose that instead of just two options, there are three. Suppose that the town can either introduce a system of one-way streets or a system of traffic lights, or it can retain the status quo. (Suppose that everyone agrees that with one-way streets, traffic lights would not be needed.) The presence of the third option changes the nature of the information on preferences that must be gathered and the manner in which it must be processed.

The Thompson Insurance Mechanism

For the Thompson insurance mechanism, one approach that will not work well for evaluating multiple options is the one suggested by Thompson, which is to consider the options one by one and discard each one that loses in a paired comparison.[16] The reason that this will not work well is that people will find it worthwhile to consider the eventual consequences of any choice and not just the value of the two alternatives being compared. Suppose for example that the status quo is to be compared with one-way streets, and whichever one wins is then to be compared with traffic lights. Consider a voter who would value one-way streets at $50 more than the status quo, and traffic lights (taking account of his share of the costs) at $100 less than the status quo. Suppose he faces the probabilistic prospects shown in Figure 16.1. That is, the official estimates, which he accepts, are that there is a 60% chance that the system of one-way streets will be approved, and if it is approved there is an 80% chance that it will then be replaced by traffic lights. On the other hand, if the system of one-way streets is not approved, then there is a 30% chance that the system of traffic lights will be approved. Starting with a system of one-way streets and no chance of the status quo, the 80% chance of traffic lights has a cost of 0.8 ($50 + $100), or $120 to this voter. Starting with the status quo and no chance of one-way streets, a 30% chance of traffic lights has a cost of 0.3 ($100) or $30. Since a certainty of one-way street is worth $50

Figure 16.1 Probabilistic prospects for three options



more than the status quo, the total expected value of moving down the left branch from the status quo is \$50 – \$120, or –\$70, while the total expected value of moving down the right branch is –\$30. Thus the voter would honestly report that moving to one-way streets would cost him \$40 compared to keeping the status quo, even though the system of one-way streets was his favorite option. If all voters had the same preferences, they would unanimously reject their favorite option.

A better way to evaluate multiple options under the Thompson insurance mechanism is to ask each voter to specify his favorite option and how much insurance he wanted to buy against getting each of the others instead of his favorite. For each option, the election authority specifies the probability that it will be chosen, with probabilities that sum to one. The voter above would specify one-way streets as his first choice and buy \$50 of insurance against the status quo and \$150 of insurance against traffic lights. The option that would be chosen would be the one against which voters had bought the least insurance. As with two options, there would be more than enough in insurance premiums to pay off the insurance claims, because the sum of the insurance premiums would be a weighted average of the amounts of insurance bought against all of the options, which must be at least as great as the minimum of the amounts of insurance. The option against which the least insurance was bought would be the most highly valued because, for each voter, the value of moving from any option, x, to any other option, y, is the difference between the amount of insurance that the voter bought against getting x and the amount of insurance that he bought against getting y. Therefore the total value of moving from x to y is the difference between the total amount of insurance bought against x and the total amount of insurance bought against y. If x is the option against which the least insurance was bought, then there is no way of moving to another option and obtaining a positive net value.

This conclusion presumes that there are no changes in the incomes of the voters. One source of changing voter incomes is the surplus that the Thompson insurance mechanism generates. If people vary with respect to the proportion of additions to income they would spend to secure their favorites among the options on which they were voting, then it would be possible for the distribution of the surplus from the Thompson insurance mechanism to change the identity of the

most valuable option. To take account of such income effects, one would need to ask each voter how his demand for insurance for each option varied with his income. Then it would be possible to identify an option that would be most valuable after the distribution of the surplus from use of the Thompson insurance mechanism.

The Demand-Revealing Process

For the demand-revealing process, income effects enter into the evaluation of multiple options in a different way. Return to the example of a choice among the status quo, one-way street and traffic lights. We might begin by asking each citizen to rank the three possibilities. Suppose a citizen reports a ranking of the status quo as his first choice, one-way streets as his second choice, and traffic lights third. We could then ask him how much he would be willing to pay to have his first choice instead of his second, and how much to have his second choice instead of his third.

Suppose that he answers "$30" to the first question and "$20" to the second. One might think that this would provide all the information that we would need from him to evaluate the options, because the value to him of having his first choice instead of his third would be the sum of the two valuations he already reported, namely $50. However, this is not generally true, because such a calculation would leave out income effects.

When the voter is trying to preserve the status quo, he is richer if the alternative to the status quo is one-way streets, his second choice, rather than traffic lights, his third choice. If he is willing to pay $30 to have the status quo rather than one-way streets and $20 to have one-way streets rather than traffic lights, he may be unwilling to pay as much as $50 to have the status quo rather than traffic lights, because the willingness to pay $30 for the status quo rather than one-way streets presumed that he had not already paid $20 to get from traffic lights to one-way streets. Having already paid the $20, he might be willing to pay only $29 rather than $30 to have the status quo rather than the one-way streets, and therefore be willing to pay only $49 to have the status quo rather than the traffic lights. Such income effects came up in the discussion of the Hicks–Kaldor criterion, in Chapter 3.

One way of interpreting such income effects is to see them as a reflection of the lack of constancy of the marginal utility of money. If it were possible to measure the changes in utility that people would experience from proposed projects, then one could use a rule that those projects and only those projects that generated net positive changes in aggregate utility would be approved. Such differences in utility would be strictly additive: The change in utility in moving from x to z would always equal the sum of the change in utility in moving from x to y and the change in utility in moving from y to z (assuming that it is the projects themselves rather than any characteristics of the process of change that generate utility for voters). But because there is not a constant factor for converting dollars

of willingness to pay into changes in a person's utility, the same additivity does not hold for willingness to pay. Any procedure for aggregating changes in utility that assigns the same marginal utility of money to a person irrespective of whether the contemplated move is from his first choice to his second, his second choice to his third, or his first choice to his third cannot be correct if the marginal utility of money changes as income changes.

When there are income effects, there is relevant information that is not contained in the relative value of adjacent options in voters' rankings; the identification of the most valuable option requires a separate determination, for each voter and for each pair of outcomes, of how much that voter is willing to pay to have the preferred outcome instead of the alternative. Thus we are faced, as we were in the case of equal weights on votes, with the need for a matrix of paired comparisons. And as with the case of equal weights on votes, there is a possibility of cycles.[17] Consider Example 16.1, of two voters evaluating three options.

Example 16.1 Two Voters Evaluating Three Options

Individual preferences

Paired comparisons

Voter	SQ	OWS	OWS	TL	SQ	TL
1	30	0	20	0	49	0
2	0	33	0	22	0	48
Total	30	33	20	22	49	48

Matrix of dollar majorities

	TL	OWS	SQ
TL	0	2	−1
OWS	−2	0	3
SQ	1	−3	0

The pattern of the matrix of majorities in Example 16.1 is the familiar pattern of cyclic preferences for three options. It appears that traffic lights are valued more highly than one-way streets, one-way streets more highly than the status quo, and the status quo more highly than traffic lights. As in the case of voting with equal weights, some device must be used to cut through such cycles. Applying the

maximin rule, the ranked pairs rule, or the Schulze method to this matrix yields the result that the system traffic lights is chosen. Applying the alternative Schwartz rule or the alternative Smith rule results in the elimination of the option of one way streets, since it has no first place votes, and then the selection of the status quo.

Then one must ask what Clarke taxes, if any, are owed by a pivotal voter. Suppose that the rule in use selects traffic lights. If Voter 2 had not been present the result would have been what Voter 1 wants, namely the status quo. And having the system of traffic lights instead of the status quo has a cost of \$49 to Voter 1. Therefore this is the Clarke tax on Voter 2. But Voter 2 places a value of just \$48 on traffic lights compared to the status quo. If Voter 2 could have anticipated this result, he would have been better off not voting. Thus this rule does not motivate voters to report their preferences honestly in all circumstances.

Looking only at Example 16.1, one might suppose that such a result could occur only when there were few voters, or that any such excess of a Clarke tax above a voter's valuation might never be more than a small fraction of a voter's valuation. But when the number of voters increases, things can get worse. Consider Example 16.2:[18]

Example 16.2 303 Voters Evaluating Three Options

Paired comparisons

Persons	SQ	OWS	SQ	TL	OWS	TL
1	0	20	0	46	0	30
1	0	30	20	0	45	0
1	47	0	30	0	0	20
100	20	0	44	0	30	0
100	30	0	0	20	0	44
100	0	44	0	30	20	0
Total	5,047	4,450	4,450	5,046	5,045	4,450

Matrix of dollar majorities

	TL	SQ	OWS
TL		596	−595
SQ	−596		597
OWS	595	−597	

Applying the maximin rule, the Condorcet rule or the ranked pairs rule to this matrix yields the result that the traffic lights are chosen. But if the first voter is deleted, the matrix of dollar majorities becomes

	TL	OWS	SQ
TL	0	−625	550
OWS	625	0	−617
SQ	−550	617	0.

The winner in this case is the status quo. Thus the first voter is pivotal, and the net cost to all other voters of having the traffic lights rather than the status quo is 550, which is about twelve times the value to the pivotal voter. In view of the possibility of such examples, it would be imperative that the Clarke tax on the pivotal voter in such cases not be computed as the net cost to all other voters.

One way of defining the Clarke tax that produces the same result as the ordinary definition when there is no cycle, but does not put voters at risk of paying more for an outcome than it is worth to them, is to generalize the principle from decisions over pairs of options, that what a pivotal voter must pay is what the voter offered to pay, minus the winning margin. Define the "generalized winning margin of *A* over *B*" as the difference between *A*'s worst loss and *B*'s worst loss. Specify that in the event of a cycle, the winner will be the determined by the maximin rule applied to the matrix of dollar (or other preference indicator) majorities, and define the "generalized Clarke tax" for votes among more than two options by the rule that if a voter causes *A* to be chosen when *B* would otherwise have been chosen, then the generalized Clarke tax for that voter is what the voter offered to pay to achieve *A* instead of *B*, minus the generalized winning margin of *A* over *B*. Note that since *A* is the maximin winner, the generalized winning margin of *A* over *B* is necessarily positive (or zero in the event of a tie), and that if there are just two options, the generalized winning margin reduces to the actual winning margin, and the generalized Clarke tax reduces to the regular Clarke tax. In the example above, the generalized winning margin of traffic lights over the status quo is −595 − (−596) = 1. Thus the generalized Clarke tax for the first voter would be 46 − 1 = 45.

Such a voting rule is not strategy proof. If the first voter stated a value of 31 instead of 30 for the difference between traffic lights and one-way streets, then the worst loss of traffic lights would fall to −594, and his Clarke tax would fall to 44.

That such strategies are possible is not surprising in view of the Gibbard–Satterthwaite theorem. The Gibbard–Satterthwaite theorem is concerned with ranking-based vote processing rules (RBVPRs). The demand-revealing process is nominally based not on rankings of outcomes, but rather on scores that specify the value to voters of having one outcome rather than another. From another perspective, however, an "outcome" can be defined as the combination of an element from the set of possibilities under consideration and a change in wealth for each voter.

Define an "outcome" as an element of the set of things that the choice process is directly concerned with (e.g., the status quo, one-way streets or traffic lights),

and an "extended outcome" as the combination of an outcome and a change in wealth for each voter.

Under the assumption that each voter is concerned only with the reduction in his or her own wealth and not with the reductions in wealth of other voters, the voters' reported willingnesses to pay to achieve one outcome instead of another provide voter rankings of the extended outcomes. However, the assumption that each voter is not concerned about changes in the wealth of other voters and prefers more wealth to less for himself, along with the use of a mechanism that makes no provision for departures from these assumptions, means that the condition of universal domain is not satisfied for the demand-revealing process conceived as a mechanism for choosing among extended outcomes. Thus the Gibbard–Satterthwaite theorem does not apply directly. Nevertheless, in view of the Gibbard–Satterthwaite theorem, it should not be surprising if any way of assigning Clarke taxes when there are more than two options yields a rule that, like all non-dictatorial RBVPRs for more than two options, is manipulable.

Based on the analysis in Chapter 9 of the conditions that make cycles likely or unlikely, it seems reasonable to expect that cycles would rarely arise with the demand-revealing process, as long as it was not used to evaluate highly redistributive options. Therefore the incentive to understate preferences that arises from the rule of charging each pivotal voter a generalized Clarke tax seems unlikely to significantly impair the efficiency of the demand-revealing process.

The income effects that generate the possibility of cycles are also responsible for another consequence that must be handled if the demand-revealing process is to be a practical voting rule. This consequence is the quadratic increase in the number of comparisons that must be made as the number of possibilities under consideration increases. When there are three possibilities, each voter must make three comparisons. When there are four possibilities, six comparisons are needed. In general, the number of comparisons needed for n possibilities is the sum of the integers from 1 to $n - 1$, that is, $n(n - 1)/2$. People could be expected to regard it as unreasonably tedious to state 45 valuations in order to describe their preferences among 10 possibilities.

One way of coping with this difficulty is to assume a mathematical structure of preferences and estimate each voter's preferences on the basis of a more limited amount of information. In particular, one might assume that each voter has a utility function of the form:

$$U = Q(m) + \ln(W_0 - P_{mx}) \tag{16.1}$$

where m is the outcome that is chosen, $Q(m)$ is the "quality" of m as perceived by the voter, W_0 is the voter's initial wealth, and P_{mx} is the payment that the voter makes to achieve m rather than x. To anchor the arbitrary quality scale, assign a quality of 0 to the voter's least favored outcome. Suppose there are four possible outcomes, a, b, c and d, and that the voter reports the following willingness to pay:

To achieve *c* instead of *d*:	\$150
To achieve *b* instead of *c*:	\$90
To achieve *a* instead of *b*:	\$40
To achieve *a* instead of *d*:	\$250

If there were no income effects, one would expect the last number to be the sum of the three preceding ones, namely \$280. The fact that the last number is less than this implies that income effects are present in the above numbers. When combined with the assumed form of the voter's utility function from (16.1), the numbers imply:

$$Q(c) + \ln(W_0 - 150) = 0 + \ln(W_0) \tag{16.2}$$

$$Q(b) + \ln(W_0 - 90) = Q(c) + \ln(W_0) \tag{16.3}$$

$$Q(a) + \ln(W_0 - 40) = Q(b) + \ln(W_0) \tag{16.4}$$

$$Q(a) + \ln(W_0 - 250) = 0 + \ln(W_0) \tag{16.5}$$

These four equations can easily be solved by successive approximations for the four unknowns, W_0, $Q(c)$, $Q(b)$ and $Q(a)$. One enters a trial value of W_0, solves (16.2) for $Q(c)$, uses the values of W_0 and $Q(c)$ to solve (16.3) for $Q(b)$, uses the accumulated information to solve (16.4) for $Q(a)$, and then compares the result with the value of $Q(a)$ obtained from (16.5). If the former value exceeds the latter, a smaller value of W_0 is needed, and if the latter exceeds the former, a larger value is needed. The process is easily generalized for any number of options, and can be accomplished by a computer in the blink of an eye. For the example, the solution is

$$W_0 = 745.87$$

$$Q(c) = 0.2245$$

$$Q(b) = 0.3531$$

$$Q(a) = 0.4082$$

These number imply in turn that the amount that the voter is willing to pay to have *a* instead of *c* is given by

$$Q(a) + \ln(W_0 - P_{ac}) = Q(c) + \ln(W_0),$$

which implies that

$$P_{ac} = \$125.17.$$

Similarly, P_{bd} can be computed to be \$221.90. Thus if one is willing to assume a particular form for a voter's utility function, the number of valuations that the voter must supply to determine his willingness to pay, to have any of n options instead of any of the $n - 1$ others, can be reduced from $n(n - 1)/2$ to n. There will be a unique solution, which a computer can readily determine, as long as the amount that the voter is willing to pay to have his most favored instead of his least favored option is between the largest amount that the voter is willing to pay to have an option instead of the one he ranks immediately below it, and the sum of all such amounts from the voter's most favored option to his second-least favored. For the example, this constraint is that the amount that the voter is willing to pay to achieve a instead of d must lie between \$150 and \$280. If any voters did not want to have the assumed functional form applied to them, they could be invited to provide all $n(n - 1)/2$ valuations, or else provide alternative algorithms by which all valuations could be computed.

Comparing the Demand-Revealing Process to the Thompson Insurance Mechanism

One might wonder why the introduction of more than two options is so much more complicated for the demand-revealing process than for the Thompson insurance mechanism. The answer is that the entitlements of voters under the Thompson insurance mechanism are mutually consistent in a way that the entitlements of voters under the demand-revealing process are not. Under the Thompson insurance mechanism, every voter has an initial entitlement to a weighted average of the outcomes, with the weights determined by the assigned probabilities. The purchases of insurance reveals which outcome provides the greatest increase in total value. The movement to that outcome is achieved as a Pareto improvement from the initial entitlements.

Under the demand-revealing process, by contrast, the entitlement of each voter is to the outcome that would occur if the voter were to abstain. Each voter is allowed to purchase departures from the consequences of abstention, at marginal cost. But for each comparison of a pair of options, the consequences of abstention are different, leading to a lack of additivity in valuations of sequences of moves.

Choosing among Options in a One-Dimensional Continuum

Examples of decisions involving options that lie in one-dimensional continuums are such things as the amount of public expenditure on parks in a city, or the speed limit on a street. While the options for such a decision may be considered to lie in a continuum, they might more accurately be described as decisions over lattice points that lie in a continuum, because expenditure on parks will be a whole

number of dollars (or pennies) and speed limits will in fact be whole multiples of 5 miles per hour. Thus it would be possible to ignore the fact that the options lie in a continuum and evaluate a set of discrete options by the method described above. But the fact that the options have a known spatial relation to each other permits additional elements in decision-making procedures.

The Thompson Insurance Mechanism

Consider a decision about how much money a town will spend on parks. The election authority would have the responsibility of specifying a probability distribution over possible outcomes. (For simplicity, suppose that it is agreed that the outcome will be spending of some multiple of $10 per voter.) From the fact that the outcomes were points on a continuum, the election authority might infer that the probabilities associated with the outcomes could be described by a function with the shape of a lognormal distribution. This would simplify the process of specifying the probabilities.

Each voter would report a most favored outcome and the amount of insurance that the voter wished to buy against every other outcome. A voter with smooth, convex preferences might reasonably approximate the loss from each alternative as a quadratic function of distance from his or her favorite. This would simplify the reporting of the voter's preferences. Apart from these simplifications in the specification of probabilities and the reporting of preferences, the operation of the Thompson insurance mechanism for options in a continuum is no different from its operation for a set of multiple discrete options.

The Demand-Revealing Process

For an examination of the operation of the demand-revealing process on options in a continuum, it is useful to consider first a decision among just three discrete options that lie in a continuum. Suppose the question is how much money will a town spend on parks, and suppose it is agreed that the budget for parks will be either $200, $300 or $400 per person. As in the case of three options that are not in a continuum, we might ask the voters to state, for each pair of options, how much they are willing to pay to achieve the one they prefer.

It is interesting to note that, with options that lie in a continuum as with option that do not, income effects can cause cycles among discrete options even when there are just two voters. Consider the preferences in Example 16.3:

Voter 1 finds the three options increasingly less attractive as the level of expenditure rises, while Voter 2 finds the options increasingly less attractive as the level of expenditure falls. Both voters have preferences that are consistent with downward sloping demand curves. For voter 2, this is reflected in the fact that the marginal value of moving from $300 to $400 (namely $26) is less than the marginal utility of moving from $200 to $300 (namely $27). The income effect induces Voter 2 to be willing to pay only $45 rather than $53 (the sum of $26 and

Example 16.3 Two Voters Evaluating Three Levels of Park Expenditure

Individual preferences

Paired comparisons

Voter	$200	$300	$300	$400	$200	$400
1	23	0	24	0	46	0
2	0	27	0	26	0	45
Total	23	27	24	26	46	45

Matrix of dollar majorities

	$200	$300	$400
$200	0	−4	1
$300	4	0	−2
$400	−1	2	0

$27) for a move from $200 to $400. For voter 1, the value of moving from $300 to $400 (namely −$24) is less than the value of moving from $200 to $300 (namely −$23), while the value of moving from $200 to $400 is just −$46 rather than −$47, because of the income effect again. Voter 1's demand curve is downward sloping, as a demand curve should be, though the values attached to increases in spending are negative.

Because of the difference in the magnitudes of the income effects, paired comparisons show that $400 beats $300, $200 beats $400, and $300 beats $200. Thus a cycle is possible with just two voters and three options in a continuum, and with preferences that behave the way demand curves normally do. If we were to apply the maximin rule proposed earlier for options that do not lie in a continuum, the winning option would be $400, because its worst loss is least bad.

Is the maximin rule a good rule for deciding the winning option among options that lie in a continuum? To address this question, it is useful to examine a case where a different rule leads to a different result. Consider Example 16.4

In Example 16.4 the preferences of the two voters for three previous options are the same as in Example 3, but now a fourth option has been added. If we use the maximin rule to select the winning option, then $400 wins again. But there is a

Example 16.4 Two Voters Evaluating Four Levels of Park Expenditure

Individual preferences as matrices of paired comparisons

	Voter 1					Voter 2			
	$200	$300	$400	$500		$200	$300	$400	$500
$200	0	−27	−45	−62	$200	0	23	46	70
$300	27	0	−26	−43	$300	−23	0	24	48
$400	45	26	0	−25	$400	−46	−24	0	25
$500	62	43	25	0	$500	−70	−48	−25	0

Matrix of dollar majorities

	$200	$300	$400	$500	Worst Loss
$200	0	−4	1	8	−4
$300	4	0	−2	5	−2
$400	−1	2	0	0	−1
$500	−8	−5	0	0	−8

possible argument that the result should be a tie between $400 and $500. The argument is that if one compares each option only with the options that are spatially beside it, then $300 beats $200 and $400 beat $300, but $500 and $400 are tied. Thus if one thinks of the best collective choice as the option such that the marginal loss by those who lose is just offset by the marginal gain by those who gain, then by this criterion $400 and $500 are tied.

However, there are two difficulties with this approach. First, the source of cycles is income effects. The demand-revealing process is an effort to add up the gains and losses of voters from selecting one outcome instead of another. Whatever measure of gains and losses is used, it will be affected by the incomes of voters, and one component of income is the option that will be chosen if one does not succeed in having one's favored option chosen. If the choice depends on income, there is nothing compelling about the incomes that yield a tie between adjoining options.

The second difficulty is that there is no obvious way that people could be motivated to report their preference truthfully if income effects were ignored. If voters were told to report only their willingness to pay for adjoining options and ignore income effects, but they were charged for the marginal costs of the their votes to others, then they would have an incentive to under-report their valuations, lest they be required to pay more for a change than it was worth. And if they were told to report income effects along with their willingness to pay for movements among adjoining options, and the decision was based on finding a tie between adjoining options, with the income effects used to set upper limits on what they

would be required to pay, then they would have an incentive to overstate their valuations of changes among adjoining options. Therefore I recommend the rule that, for choices among options on a continuum, the outcome should be selected by the maximin rule. For Example 16.4, this would result is the selection of $400 as the level of per capita expenditure.

Next consider the question of whether cycles are likely for options on a continuum. The answer is that cycles are almost inevitable if one uses a fine enough lattice, but they may nevertheless be of no practical consequence.

Suppose that one uses a lattice so fine that there are many lattice points between any two points that are the first choices of different voters. With such a lattice, find a pair of adjacent lattice points that tie with each other, or come as close or closer to tying than any other pair. Call the left point of this pair of lattice points C. As one considers points between C and the first point to the left of C that is the first choice of some voter, there is one constant set of voters, L, who prefer movement to the left and another constant set, R, who favor movement to the right.

Income effects tend to be proportional to the change in income, so assume that the income effect on the value of choices among lattice points for the set L is $k_L \Delta I$, where ΔI is the change in income that induces the income effect. Similarly, $k_R \Delta I$ is the income effect for the set R. It would be a coincidence of essentially zero probability if k_L were equal to k_R, so without loss of generality, assume that k_L is less than k_R.

Consider C and the two lattice points just to the left of it, A and B, as in Figure 16.2. Since these lattice points are to the left of where the tie or near–tie occurs, the aggregate value of a move from A to B is positive, as is a move from B to C.

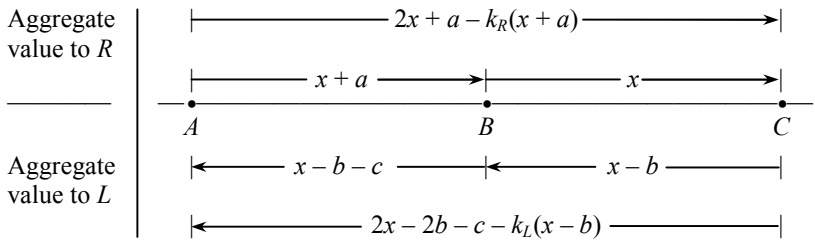


Figure 16.2 Valuations of movements among three lattice points

Now consider a move from C to A. The expressions in the top half of Figure 16.2 specify the amounts of money that the set of voters R is willing to pay to achieve an outcome instead of an alternative to the left. The expressions in the bottom half are the amounts of money that set L is willing to pay to achieve an outcome instead of an alternative to the right. The explanations of the amounts are as follows.

The amount of money set R is willing to pay to have C instead of B is defined as x. Since this set wishes to move the outcome to the right and since the cost of additional departures of a given size tends to increase with distance from one's ideal, the amount that they are willing to pay to have B instead of A is slightly more, namely $x + a$. When set R is thinking about getting to outcome C from A rather than from B, they are poorer by $x + a$. Therefore the amount of money that they are willing to pay to have C rather than A is the sum of the amounts they are willing to pay to have B rather than A and C rather than B, minus $k_R(x + a)$.

Now consider set L. Since point C is slightly to the left of the tie or near tie, the amount that set L is willing to pay to have B instead of C is less than what set R is willing to pay to have C instead of B, by an amount that is specified as b. Since set L has downward sloping demand curves, the amount that they are willing to pay to have A instead of B is less than what they are willing to pay to have B instead of C, by an amount that is specified as c. Because of the income effect, the amount that they are willing to pay to have A instead of C is less than the sum of these two amounts by the product of k_L and the amount they are willing to pay to have B instead of C.

Summing the valuations of the two groups, the aggregate value of a move from A to C is $a + 2b + c + k_L(x - b) - k_R(x + a)$. In the limit as the distance between lattice points approaches zero, the ratios of a, b and c to x approach zero. Therefore, in the limit as the distance between lattice points approaches zero, the aggregate value of a move from A to C approaches $(k_L - k_R)x$. Since k_R is greater than k_L, this is negative, implying that a move from C to A has a positive aggregate value. Thus there is a cycle. The aggregate value of moves from A to B, from B to C, and from C to A are all positive.

If k_R had been less than k_L, then a cycle would have been found to the right of the tie or near-tie. Cycles can be completely avoided only if k_R is equal to k_L. Thus it is practically certain that cycles can be found if choices in a one-dimensional continuum are examined as a sufficiently fine lattice. Such cycles, however, do not make the demand-revealing process unworkable. It is still possible to use the maximin rule (the winning option is the one whose worst loss is least bad) to determine the outcome, combined with generalized specification of Clarke taxes that was developed for multiple discrete options.

The principle advantage of knowing that options lie in a continuum is that it is then possible to structure reports of voters' preferences among options based on the way preferences tend to vary with the spatial locations of options. The number of dollars that a person is willing to pay per unit of change in a parameter to be decided has the character of a demand schedule. Thus, rather than specifying a willingness to pay to have options rather than less preferred ones, a voter can specify a demand schedule for changes in the outcome, that is, a willingness to pay for unit changes in the outcome, as a function of location on the continuum.

Returning to the example of spending on a park, a voter might vote by answering a series of questions like the following:

1. Taking $100 to $400 per person as the range of possible park spending, and recognizing that you are going to be required to pay your share of whatever spending level is decided, what level of spending would you find most attractive?
2. If, when the preferences of all voters are aggregated, it turns out that the tentative outcome is different from your first choice, are you willing to pay anything to have a final outcome that is closer to your first choice? (If not, you do not need to answer further questions)
3. If the tentative outcome was $10 per person less than your first choice, how much would you be willing to pay to get your first choice instead? Would the answer be any different if the tentative outcome was $10 more than your first choice?
4. Suppose the tentative outcome was $50 less than your first choice. How much would you be willing to pay to raise it by $10? By $20? If it was $50 more than your first choice, how much would you be willing to pay to reduce it by $10? By $20?
5. Suppose the tentative outcome was $100 less than your first choice. How much would you be willing to pay to raise it by $10? By $20? If it was $100 more than your first choice, how much would you be willing to pay to reduce it by $10? By $20?
6. What is the maximum amount of money that you would want to spend to influence the outcome?

With such a series of questions, it is possible to estimate a voter's willingness to pay with respect to all pairs of options. After this has been done for all voters, one can form the aggregate matrix of paired comparisons and identify the option that wins by the maximin rule. Then the generalized Clarke taxes can be calculated for all voters.

For a voter who behaved as Margolis predicts and expressed a willingness to pay to achieve what the voter believed to be socially right rather than what was personally advantageous, the answers would reveal a relatively stable willingness to pay for a move in the direction of the voter's first choice, irrespective of the starting point or the magnitude of the move. The rule would have no difficulty accommodating such votes.

Other Processes for Options in a Continuum

Dreze and de la Vallee Poussin proposed that decisions about options in a continuum be made by having each voter specify an amount of money per unit of the continuum that the voter would be willing to pay to move in the direction that the voter preferred, an amount that would also serve as compensation per unit of movement in the opposite direction. Movement would then occur in the direction with greater aggregate value, and everyone who reported true preferences would be better off.[19] Malinvaud proposed the same idea independently.[20] This idea is

essentially a simplified version of the Thompson insurance mechanism, with the probability of an increase restricted to equal the probability of a decrease. Any voter who believed that a movement in one direction was more likely than a movement in the other direction would have an incentive to report that he would require substantial compensation for movement in the likely direction. Such behavior could be expected to make such a decision procedure unworkable.

I proposed a variation on this mechanism with a suggestion for overcoming the strategy problem.[21] The decision process would start with the election authority estimating the efficient outcome. It would also estimate the elasticity of demand. As in the Dreze–de la Vallee–Malinvaud mechanism, voters would report amounts that they would be willing to pay for incremental movement in the preferred direction or accept as compensation for incremental movements in the opposite direction. But now, instead of implementing an incremental movement in the more highly valued direction, the election authority would re-estimate the efficient outcome, using the information provided by voters. If voters believed that the selected outcome was an unbiased estimate of the efficient outcome, they would have no incentive to misstate their preferences, and even if they thought the estimate was slightly inaccurate, risk aversion might induce them to report their preferences accurately.

Conclusion

The Dreze–de la Vallee–Malinvaud procedure is unworkable because it offers incentives for misstatements whenever voters do not believe that increases and decreases in the parameter to be determined are equally likely. The related procedure that I developed and the Thompson insurance mechanism share the limitation that voters have incentives to misstate their preferences unless they believe that the election authority has made an accurate estimate of obscure information. In the case of my mechanism, the obscure information is the expected value of the final decision. In the case of the Thompson insurance mechanism, the obscure information is the probability that a proposal will be approved. While these mechanisms work in theory when the election authority persuades voters that the authority has made the right estimates, I believe that in practice they would be neither efficient nor fair. They would not be efficient because people who believed the official estimates were wrong would have an incentive to give misinformation. If a person did not care whether one-way street were implemented and believed that the probability that they would be implemented was 30% when the official probability was 25%, then he would have an incentive to claim a false loss from their implementation, with the magnitude of that claim limited only by his aversion to risk. The difficulty of deterring such false claims would make it impossible to achieve efficiency by using Thompson's suggestion.

If this were the only difficulty, one might say that if false reports precluded efficiency, at least everyone who reported their losses truthfully was compensated. But such a perspective ignores the redistribution involved in considering proposals. If there is a 25% chance that a system of one-way streets will be implemented, then a person who would experience a loss of $400 from their implementation has a $100 loss of insurance premiums from the fact that the proposal is considered. This makes it very important that the assigned probability be "correct." But there is no objectively "correct" probability. After the fact, the proposal was either implemented or it was not. Everyone whose favored outcome is eventually approved can reasonably complain that if a higher probability had been assigned to the eventual outcome, then they would not have had to pay as much. I believe that these difficulties make mechanisms based on uncertainty in the outcome unworkable. Thus I believe any workable vote-processing rule with endogenous weights will be based, like the demand-revealing process, on marginal cost pricing.

Notes

1 For a discussion of the issue, see *Journal of Economic Perspectives* 8/4 (1994).
2 Earl A. Thompson, "A Pareto Optimal Group Decision Process," in Gordon Tullock (ed.), *Papers on Non-Market Decision Making*, (Charlottesville, 1966), pp. 133–40.
3 Martin Bailey, *Constitution for a Future Country* (New York, 2001), 182–8.
4 Thompson, "A Pareto Optimal Group Decision Process," p. 137.
5 Bailey, *Constitution for a Future Country*, pp. 40–43.
6 Edward H. Clarke, "Multipart Pricing of Public Goods: An Example," in Selma Mushkin (ed.), *Public Prices for Public Products*, (Washington, 1972), pp. 125–30. An earlier article of Clarke's, "Multipart pricing of Public Goods," *Public Choice* 11 (1971): 17–33, dealt with decisions about continuous parameters.
7 T. Nicolaus Tideman and Gordon Tullock. "A New and Superior Process for Making Social Choices," *Journal of Political Economy* 84 (1976): 1145–60.
8 Howard Margolis, "A Note on Demand-Revealing," *Public Choice* 40 (1983): 217–26.
9 Bailey, *Constitution for a Future Country*, pp. 192–4, 204–30.
10 For fuller discussions of this issue, see I. J. Good, "Justice in Voting by Demand Revelation," *Public Choice* 29/2 (1977): 65–70, and T. Nicolaus Tideman, "Efficient Local Public Goods Without Compulsory Taxes," *Perspectives on Local Public Finance and Public Policy* 2 (1985): 197–9.
11 Bailey, *Constitution for a Future Country*, pp. 82–5.
12 For a discussion of this evidence, see Shabnam Modarres Mousavi, "Methodological Foundations for Bounded Rationality as a Primary Framework," doctoral dissertation, Economics Department, Virginia Polytechnic Institute and State University (2002), p. 59.
13 T. Nicolaus Tideman, "An Experiment in the Demand-Revealing Process," *Public Choice* 41 (1983): 387–401.
14 Martin Bailey, *Constitution for a Future Country*, pp. 82–5, 173–6.

15 J.E. Calfee and P.H. Rubin, "Some Implications of Damage Payments for Nonpecuniary Losses," *Journal of Legal Studies* 21 (1992): 371–411.
16 Thompson, "A Pareto Optimal Group Decision Process," p. 135.
17 This possibility was pointed out by Tibor Scitovsky in "A Note on Welfare Propositions in Economics," *Review of Economic Studies* 9 (1941): 77–88.
18 I am indebted to Daniel Richardson for pointing out to me the possibility of such an example.
19 J.H. Dreze and D. de la Vallee Poussin. "A Tâtonnement Process for Public Goods," *Review of Economic Studies* 38 (1971): 133–50.
20 E. Malinvaud, "A Planning Approach to the Public Good Problem," *Swedish Journal of Economics*, 73/1 (1971): 96–112.
21 T. Nicolaus Tideman, "The Efficient Provision of Public Goods," in Selma Mushkin, (ed.), *Public Prices for Public Products*, (Washington), 1972, 111–23.

This page intentionally left blank

Chapter 17
Lessons from the Excursion

What is to be learned from this excursion through the territory of collective decisions and voting? This final chapter summarizes the main points of the preceding chapters.

Part I of the book discusses collective decisions in general. A decision is not an action, but rather an orientation of intentions. A collective decision is a process that orients the intentions of the members of a collectivity to a pattern of coordinated future actions, so that in the ordinary course of events, the pattern is realized.

Collective decisions can be made in a variety of ways, either by agreeing on an outcome directly or by first agreeing on a procedure and then applying the procedure. Any collectivity's first decision to apply a procedure is necessarily made by some direct method.

Direct ways of making collective decisions vary in the dimensions of compensation, dissatisfaction, and effort. When neither compensation nor dissatisfaction is present, the mode is consensus. Introduce compensation, and the mode becomes trade. If dissatisfaction is present and compensation is not, the mode is pseudo-consensus. If both compensation and dissatisfaction are present, the mode is extortion.

Procedures that people might agree to use for collective decisions fall into the categories of authority, random process, voting and contest. Many collective decisions combine elements from different categories.

Economists have proposed a variety of criteria for identifying good collective decisions. Pareto suggested that if a change makes some persons better off and others worse off, then there is no way of knowing whether the change is good overall, so only a change that makes no one worse off can properly be labeled an improvement. This criterion presumes implicitly that it is possible to specify the initial position (the path of events into the indefinite future in the absence of a decision) and the alternative that will prevail if a decision is made. It also presumes implicitly that the initial position is ethically worthy of endorsement.

The Bergson–Samuelson social welfare function codifies the ethical judgments of one individual, but it provides no basis for identifying an individual whose judgments deserve to be codified. The Hicks–Kaldor criterion can be regarded as a special case of the Bergson–Samuelson social welfare function, assigning the same importance to a dollar of value to all persons at all times. While this is one way of cutting through the Bergson–Samuelson problem of identifying the source of value judgments, it leads, as Scitovsky pointed out, to contradictions.

Under the Buchanan–Tullock–Rawls criterion, people seek rules that all would agree on if they did not know their particular circumstances. This is a useful way of structuring the effort to find agreement. However, there is a danger that this practice will degenerate into the autocratic, "I am justified in imposing this rule on you, because I am sure that you would agree to it if you did not know your particular circumstances."

For the purpose of comparing ways of making collective decisions, these general perspectives from economics are not as useful as some direct criteria. The direct criteria can be divided into three broad categories: efficiency, equity and stability. Efficiency can be divided into procedural efficiency and outcome efficiency. Equity can be divided into symbolic equality, material equality and productivity-based inequality. Stability can be divided into restoration and preservation.

When the modes of collective decision-making are compared by these criteria, consensus comes out very well in terms of outcome efficiency, symbolic equality, material equality, and preservation. Pseudo-consensus is sometimes better than nothing when consensus appears to be unattainable. Trade excels in productivity-based inequality. Extortion is recommended only by the idea that it may be better than the alternative of no coordination, resulting in assault on the victim by the extorter.

The mode of authority excels in terms of procedural efficiency. Random process has high scores in both procedural efficiency and symbolic equality. Voting is as good as random process in terms of symbolic equality, and is better by every other criterion except procedural efficiency. Contest is better than voting with respect to productivity-based inequality and possibly with respect to outcome efficiency, though probably worse by the other criteria.

The next question that arises is, how do the criteria for good collective decision rules relate to the question of what is a good decision? Eschewing frameworks that seek a basis for imposing on people outcomes or rules that they wish to reject, I suggest that a good collective decision is one that 1) is satisfying ex ante to those who make the decision; 2) does not intrude on those outside the collectivity that makes the decision; and 3) provides an adequate exit opportunity for anyone in the collectivity who is unhappy with the decision.

Part II of the book deals with collective decisions by the mode of voting. A voting procedure has five elements: setting the agenda, eliciting evaluations, processing the evaluations, dealing with ties, and inducing intentions by an authoritative announcement of the result.

There are four critical dimensions in which voting situations can differ: whether the voters are presumed to be disinterested judges or self-interested advocates; the number and configuration of the options; the relative weights of the votes; and the number of options to be selected.

The most traditional voting situation involves selecting one of two options, with equal weights on the votes and ambiguity as to whether the voters are

disinterested judges or self-interested advocates. "Majority rule" for such situations seems to have evolved separately in Greek and German cultures. It can be defended as the rule that, with limited information, chooses the option that is most likely to be right (if the voters are disinterested judges) or best (if the voters are self-interested advocates). Weighting the votes can produce better results if some identifiable voters are more likely than others to be good judges or if some identifiable voters are likely to be more intensely affected than others. In modern voting theory, majority rule is characterized by necessary and sufficient conditions.

When the number of options under consideration goes from two to three or more, a possibility of cycles in majority rule arises. This possibility makes it clear, if it were not otherwise, that the statement "a majority prefers *A* to *B*" cannot be used to deduce that *A* is better than *B*. "Is better than" is inherently transitive, but preference by a majority is not. When there are cycles and decisions are made by majority rule over pairs of options, the outcome depends on the strategic and bargaining skills of the voters.

When the options lie in a single dimension in which voters have single-peaked preferences, cycles cannot arise, but when there is a continuum of options of two or more dimensions, cycles are nearly inevitable if one considers options that are arbitrarily close to one another. In real voting situations, the number of options is generally fairly small, and shared perceptions generate patterns in the preferences of voters that make a cycle involving top candidates relatively rare in practice, although cycles are ubiquitous when redistributive question are under consideration.

Arrow's famous and widely misunderstood theorem is derived from the possibility of cycles. The true import of this theorem is that under any reasonable voting rule, it is possible that the question of whether *A* or *B* is regarded as the best available option will depend on whether *C* is also under consideration. The Gibbard–Satterthwaite theorem shows that under any reasonable voting rule, it will not be possible to motivate voters in all circumstances to report their preferences honestly.

Recognizing that all voting rules have limitations, it is reasonable to evaluate voting rules in terms of properties that fall into four categories: consistency, responsiveness, stability, and qualitative attractiveness. Consistency contains the properties of majority consistency, mutual majority consistency, Condorcet consistency, Smith consistency, Schwartz consistency, and invariant loss consistency. The responsiveness category contains the properties of positive responsiveness, resolvability, and perfect-tie responsiveness. Stability contains the properties of monotonicity, homogeneity, combinativity, and independence of clones. Qualitative attractiveness includes lucidity, ease of use, cost of computation, and resistance to strategy. While Condorcet consistency, perfect-tie responsiveness and combinativity and all seem like essential properties, no voting rule can possess all three.

When voting rules that have been proposed are compared by these criteria, their acceptability depends on the degree to which complexity can be tolerated. Thirteen examined rules are arguably unacceptable under any circumstance. Plurality, approval voting and two-ballot majority rule are all acceptable if a very high value is placed on simplicity. If it is permissible to ask voters to rank candidates but not permissible to undertake sophisticated vote-counting procedures that use a matrix of majorities, then the alternative vote is attractive. If ranking and sophisticated processing of the votes are both permissible, then the maximin rule, the ranked pairs rule, and the Schulze method are attractive if one believes that monotonicity is essential; the alternative Schwartz and alternative Smith rules are attractive if one is willing to sacrifice monotonicity to achieve the highest possible resistance to strategy.

When the purpose of voting is to select an option from a one-dimensional continuum, and it is reasonable to presume that voters have single-peaked preferences in that dimension, then it may be attractive to ask each voter to report his or her first choice and select the median of these, since this will be a majority-rule equilibrium.

If a choice is to be made from a continuum of more than one dimension and voters can be presumed to have preferences in each dimension that are independent of the choice(s) in the other dimension(s), or if simplicity in the voting rule is very important, then it is sensible to ask for preferences in each dimension separately and pick the median of the voters' first choices in each dimension. If the dimensions are measurable in the same metric and a little complexity is tolerable, then it may be attractive to use the minimum sum of distances rule, thereby gaining independence of rotations.

For decisions in multidimensional continuums, when preferences in each dimension are not independent of choices in other dimensions and simplicity for users is moderately important, it may be attractive to use the multistage minimum sum of forces rule, when the dimensions have the same metric, and to use the multistage multidimensional balanced pulls method when there is no common metric for the dimensions. If simplicity for uses is not at all important, then one might ask voters to specify their preferences at all possible outcomes and use the minimum sum of forces method when dimension have a common metric, and the multidimensional balanced pulls method when they do not.

When options are discrete and more than one is to be chosen, the most attractive voting rule depends on the purpose of the decision. If the purpose is to select the candidates that seem to be best, then it is sensible to use a voting rule that ranks the candidates and select the top-ranked candidates. However, if the purpose is to select a set of candidates that represent the voters, then a different type of voting rule is needed. Party list systems are unattractive for the power that they give to the party officials that determine the lists. Cumulative voting is unattractive for the organizing costs that it imposes on like-minded voters to make their votes most effective. Some form of the single transferable vote (SVT) is

more attractive. Which version is most attractive depends on tolerance for complexity. If computers are available to do the counting, then a version like Meek's has important advantages over earlier hand-counting methods. The question of whether Meek's method or Warren's method for dealing with surpluses remains controversial among the most knowledgeable advocates of STV. With further research, it may be possible to recommend some version of the controlled pairing of outcomes by STV, which has the advantage of employing a global criterion rather than sequential eliminations. When many voters have only incomplete preferences, it may be attractive to allow voters to specify a party list as their preference after they have run out of candidates that they wish to rank individually.

In some voting situations voters may want to use decision rules that will be responsive to the intensities of voters' preferences. Two type of voting rules have been proposed for this purpose, ones that rely on uncertainty in the outcome and seek to motivate voters to report their preferences honestly by providing opportunities to insure against disfavored outcomes, and ones that rely on marginal cost pricing. Among the rules that rely on uncertainty in the outcome, the insurance mechanism proposed by Thompson is the most attractive. This mechanism gives voters entitlements to weighted averages of possible outcomes, with the weights determined by assigned probabilities that the outcomes will be chosen. The difficulty of identifying probabilities that would not be regarded as arbitrary makes it unlikely that such a mechanism would be accepted as fair.

The mechanism based on marginal cost pricing is the demand-revealing process. The most serious difficulty with this mechanism is its susceptibility to coalitions. Incorporating compensation into options, to make them as non-redistributive as is reasonably feasible, mitigates this problem. Having a voting procedure that does not permit voters to get written evidence of how they voted also reduces the problem of coalitions. When there are more than two options, income effects can create cycles with the demand-revealing process. It is possible to cut through any such cycles using the same procedures that are used with rules that are not responsive to the intensities of preferences. Cycles remove the absolute guarantee of incentives for truthful reporting of preferences that characterizes decisions between pairs of options, but this is not a serious practical problem.

To sum it all up, there are a wide variety of advanced voting rules that might be attractive to people who examined them further.

This page intentionally left blank

Bibliography

Ackerman, Bruce, *Social Justice in the Liberal State* (New Haven: Yale University Press, 1980).

Arrow, Kenneth, *Social Choice and Individual Values* (first edition, New York: John Wiley and Sons, 1951; second edition, New Haven: Yale University Press, 1963).

Bailey, Martin, *Constitution for a Future Country* (New York: Palgrave, 2001).

Balinski, Michael and Young, H. Peyton, *Fair Representation: Meeting the Ideal of One Man, One Vote* (New Haven: Yale University Press, 1982).

Bergson, Abram, (writing as Burke), "A Reformulation of Certain Aspects of Welfare Economics," *Quarterly Journal of Economics*, 52 (1938): 310–34.

Bernholz, Peter, "Logrolling, Arrow Paradox and Cyclical Majorities," *Public Choice*, 15 (1973): 87–95.

Black, Duncan, "On the Rationale of Group Decision-Making," *Journal of Political Economy*, 56 (1948): 23–34.

———, *The Theory of Committees and Elections* (Cambridge: Cambridge University Press, 1958).

Blau, Julian, "The Existence of Social Welfare Functions," *Econometrica*, 25 (1957): 302–13.

Blaug, Mark, *Economic Theory in Retrospect* (Homewood, Illinois: Richard D. Irwin, 1962).

Bordes, Georges and Tideman, Nicolaus, "Independence of Irrelevant Alternatives in the Theory of Voting," *Theory and Decision*, 30 (1991): 163–86.

Bowen, Howard R., "The Interpretation of Voting in the Allocation of Economic Resources," *Quarterly Journal of Economics*, 58 (1943): 27–48.

Brams Steven J., and Fishburn, Peter C., *Approval Voting* (Boston: Birkhäuser, 1983).

Browning, Edgar K., "A Note on Cyclical Majorities," *Public Choice*, 12 (1972): 111.

Buchanan, James, "Social Choice, Democracy, and Free Markets," *Journal of Political Economy*, 62 (1954): 114–23.

———, *The Limits of Liberty: Between Anarchy and Leviathan* (Chicago: University of Chicago Press, 1975).

Buchanan, James and Tullock, Gordon, *The Calculus of Consent* (Ann Arbor: University of Michigan Press, 1962).

Calfee, J.E., and Rubin, P.H., "Some Implications of Damage Payments for Nonpecuniary Losses," *Journal of Legal Studies*, 21 (1992): 371–411.

Carroll, Lewis, *Through the Looking Glass, and What Alice Found There* (New York: Ariel Books/Alfred A. Knopf, 1986).

Clarke, Edward H., "Multipart Pricing of Public Goods," *Public Choice*, 11 (1971): 17–33.

———, "Multipart Pricing of Public Goods: An Example," in Selma Mushkin (ed.), *Public Prices for Public Products* (Washington: The Urban Institute, 1972), pp. 125–30.

Condorcet, Marie Jean Antoine Nicolas de Caritat, Marquis de, *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix* (Paris, 1785).

Coombs, Clyde, *A Theory of Data* (New York: Wiley, 1964).

Cretney, Blake, Web site Condorcet.org (http://condorcet.org).

Dreze, J.H., and de la Vallee Poussin, D., "A Tâtonnement Process for Public Goods," *Review of Economic Studies*, 38 (1971): 133–50.

Dummett, Michael, *Voting Procedures* (Oxford: Clarendon Press, 1984).

Elgie, Robert, "Two-Ballot Majority Electoral Systems," *Representation*, 34 (1997): 89–94.

Fishburn, Peter C., *The Theory of Social Choice* (Princeton: Princeton University Press, 1973).

Gibbard, Allan, "Manipulation of Voting Schemes: A General Result," *Econometrica*, 41 (1973): 587–601.

Good, I.J., "Justice in Voting by Demand Revelation," *Public Choice*, 29-2 (1977): 65–70.

Good, I.J. and Tideman, T.N., "From Individual to Collective Ordering through Multidimensional Attribute Space," *Proceedings of the Royal Society of London, A*, 347 (1976): 371–85.

———, "The Relevance of Imaginary Alternatives," *Journal of Statistics and Computer Simulation*, 12 (1981): 313–15.

Goodman, Leo, "On Methods of Amalgamation," in R.M. Thrall, C.H. Coombs, and R.L. Davis (eds), *Decision Processes* (New York: Wiley, 1954).

Green, Jerry and Laffont, Jean-Jacques, *Incentives in Public Decision-Making* (Amsterdam: North-Holland, 1979), pp. 16–19.

Green-Armytage, James, "Cardinal-Weighted Pairwise Comparison," *Voting Matters*, 19 (2004): 6–13 (www.mcdougall.org.uk/VM/MAIN.HTM).

Greenberg, Joseph, "Consistent Majority Rules Over Compact Sets of Alternatives," *Econometrica*, 47 (1979): 627–36.

Heinberg, John G., "Theories of Majority Rule," *American Political Science Review*, 26 (1932): 452–69.

Hicks, J.R., "Foundations of Welfare Economics," *Economic Journal*, 49 (1939): 696–712.

Hill, I.D., "Some Aspects of Elections—to Fill One Seat or Many," *Journal of the Royal Statistical Society, Series A* (General), 151, part 2 (1988): 243–75.

Hill, I.D., Wichmann, B.A., and Woodall, D.R., "The Single Transferable Vote by Meek's Method," *Computing Journal*, 30 (1987): 277–81.

Hoag, Clarence and Hallett, George, *Proportional Representation* (New York: Macmillan, 1926).

Hobbes, Thomas, *Leviathan* (Indianapolis: The Bobbs-Merrill Company, 1958).

Hylland, Aanund and Zeckhauser, Richard, "A Mechanism for Selecting Public Goods when Preferences Must be Elicited," Kennedy School of Government Discussion Paper 70D, Harvard University, August 1979.

Journal of Economic Perspectives, 8/4 (1994).

Kaldor, N., "Welfare Propositions in Economics," *Economic Journal*, 49 (1939): 549–52.

Kelly, Jerry S., *Arrow Impossibility Theorems* (New York: Academic Press, 1978).

Kemeny, John, "Mathematics without Numbers," *Daedalus*, 88 (1959): 571–91.

Lakeman, Enid, *How Democracies Vote: A Study of Electoral Systems* (London: Faber and Faber, 1974).

Larsen, J.A.O., "The Origin and Significance of the Counting of Votes," *Classical Philology*, 44 (1949): 164–81.

Lea, Henry Charles, *The Duel and the Oath* (Philadelphia: University of Pennsylvania Press, 1974).

Lindgren, B.W., *Statistical Theory* (New York: The Macmillan Company, 1962).

Locke, John, *Two Treatises of Government* (New York: Cambridge University Press, 1960).

Malinvaud, E., "A Planning Approach to the Public Good Problem," *Swedish Journal of Economics*, 73/1 (1971): 96–112.

Margolis, Howard, "A Note on Demand-Revealing," *Public Choice*, 40 (1983): 217–26.

May, K.O., "A Set of Independent Necessary and Sufficient Conditions for Simple Majority Decision," *Econometrica*, 20 (1952): 680–84.

McKelvey, Richard D., "Intransitivities in Multidimensional Voting Models and Some Implications for Agenda Control," *Journal of Economic Theory*, 12 (1976): 472–82.

MacLean, Iain and Urken, Arnold B., *Classics of Social Choice*, (Ann Arbor: University of Michigan Press, 1995).

Meek, Brian, "Une nouvelle approache du scrutin transférable," *Mathématiques et Sciences Humaines* 25 (1969): 13–23.

Mousavi, Shabnam Modarres, "Methodological Foundations for Bounded Rationality as a Primary Framework," Doctoral dissertation, Virginia Polytechnic Institute and State University (2002).

Newland, Robert and Britton, Frank, *How to Conduct an Election by the Single Transferable Vote*, London: Electoral Reform Society of Great Britain and Ireland, 1976.

Nozick, Robert, *Anarchy, State, and Utopia* (New York: Basic Books, 1974).

————, *Philosophical Explanations* (Oxford: Oxford University Press, 1981).

Pareto, Vilfredo, *Manual of Political Economy* (New York: Augustus M. Kelly, 1971 [1906]).

Plott, Charles, "A Notion of Equilibrium and Its Possibility under Majority Rule," *American Economic Review*, 57 (1967): 787–806.

Polinsky, A. Mitchell, "Probabilistic Compensation Criteria," *Quarterly Journal of Economics*, 86 (1972): 407–25.

Raifa, Howard and Schlaifer, Robert, *Applied Statistical Decision Theory* (Cambridge: The MIT Press, 1961).

Rawls, John, "Justice as Fairness," *Philosophical Review*, 67/2 (1958): 164–94.

Rumbaugh, Duane M. and Savage-Rumbaugh, Sue, "Apes and Language Research," in Harlan B. Miller and William H. Williams (eds), *Ethics and Animals* (Clifton, New Jersey: Humana Press, 1983), pp. 213–15.

Samuelson, Paul, *Foundations of Economic Analysis* (Cambridge, Mass.: Harvard University Press, 1947).

Satterthwaite, Mark, "Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions," *Journal of Economic Theory*, 10 (1975): 187–217.

Schmeidler, D. and Sonnenshein, H., "The Possibility of a Cheat-Proof Social Choice Function: A Theorem of A. Gibbard and M. Satterthwaite," Discussion paper No. 89, Economics Department, Northwestern University, 1974.

Schulze, M., "A New Monotonic and Clone-Independent Single-Winner Election Method," *Voting Matters*, 17 (2003): 9–19.
(www.mcdougall.org.uk/VM/MAIN.HTM).

Schwartz, Thomas, *The Logic of Collective Choice* (New York: Columbia University Press, 1986).

Scitovsky, Tibor, "A Note on Welfare Propositions in Economics," *Review of Economic Studies*, 8 (1941): 77–88.

Sen, Amartya, "Social Choice Theory: A Re-examination," *Econometrica*, 45 (1977): 53–89.

Smith, John H., "Aggregation of Preferences with Variable Electorate," *Econometrica*, 41/6 (1973): 1027-41.

Stafford, Elizabeth A., "Classic Writings in Social Choice, Fiscal Theory and Policy and State and Local Government Finance," Master's thesis, Virginia Polytechnic Institute and State University (1980).

Suzumura, Kotaro *Rational Choice, Collective Decisions and Social Welfare* (Cambridge: Cambridge University Press, 1983).

Terrell, George, *Mathematical Statistics* (New York: Springer, 1999).

Thompson, Earl A., "A Pareto Optimal Group Decision Process," in Gordon Tullock (ed.), *Papers on Non-Market Decision Making* (Charlottesville: University of Virginia, 1966), pp. 133–40.

Tideman T. Nicolaus, "Land Use Control through Administered Compensation," *Proceedings of the Meetings of the American Real Estate and Urban Economics Association* (1969): 287–93.

———, "Efficient Provision of Public Goods" in S. Mushkin (ed.), *Public Prices for Public Products* (Washington: The Urban Institute, 1972), pp. 111–23.

———, "An Experiment in the Demand-Revealing Process," *Public Choice*, 41 (1983): 387–401.

———, "Efficient Local Public Goods without Compulsory Taxes," *Perspectives on Local Public Finance and Public Policy*, 2 (1985): 181–202.

———, "A Majority-Rule Characterization with Multiple Extensions," *Social Choice and Welfare*, 3 (1986): 17–30.

———, "Independence of Clones as a Criterion for Voting Rules," *Social Choice and Welfare*, 4 (1987): 185–206.

———, "Commons and Commonwealths: A New Framework for the Justification of Territorial Claims," in Robert Andelson (ed.) *Commons Without Tragedy* (London: Shepheard-Walwyn, 1991), pp. 109-29.

———, "The Single Transferable Vote," *Journal of Economic Perspectives*, 9/1 (1995): 27–38.

Tideman, T. Nicolaus and Tullock, Gordon, "A New and Superior Process for Making Social Choices." *Journal of Political Economy*, 84 (1976): 1145–59.

Tuleja, Tad *Fabulous Fallacies* (New York: Harmony Books, 1982).

Tullock, Gordon, *Toward a Mathematics of Politics* (Ann Arbor: University of Michigan Press, 1967).

———, "Why So Much Stability?" *Public Choice*, 37 (1981): 189–202.

Vickrey, William, "Utility, Strategy, and Social Decision Rules," *Quarterly Journal of Economics*, 74 (1960): 507–35.

Warren, C.H.E., "Counting in STV Elections," Unpublished paper (1983), 19 Rectory Road, Farnborough, Hants, GU14 7BU, England.

Wichman, Brian, computer program that counts votes by "Tideman's method" (1988), available from Wichmann at 5 Ellis Farm Close, Woking, GU22 9NQ, England.

Young, H. Peyton., "Extending Condorcet's Rule," *Journal of Economic Theory*, 16 (1977): 335–53.

———, "Condorcet's Theory of Voting," *American Political Science Review*, 82 (1988): 1231–44.

Zavist, Thomas M. and Tideman, T. Nicolaus, "Complete Independence of Clones in the Ranked Pairs Rule," *Social Choice and Welfare*, 6 (1989): 167–73.

Zeckhauser, Richard, "Voting Systems, Honest Preferences and Pareto Optimality," *American Political Science Review*, 67 (1973): 934–46.

This page intentionally left blank

Index

agreement on outcome 10–15
agreement on procedure 15–18
allocation of fractions under proportional representation
  Jefferson's method 267
  Webster's method 266–7
alternative Schwartz rule 232–5
alternative Smith rule 232–5
alternative vote 193–5
Andrae, Carl 270
anonymity
  as a condition for RBVPRs 149–50
  in majority rule 85
approval voting 170–4
Arrow Theorem 123–37
authority as a mode 16, 54, 66–7
Bailey, Martin xiv–xv, 290–92, 295, 298, 300
Bergson, Abram 26–8
Black, Duncan 122 fn, 178, 195, 209, 213–15
Black rule 209–11
Blaug, Mark 23
Borda, Jean-Charles de 176–7
Borda rule 176–81
Bordes, Georges xiii, 142 fn
Buchanan, James 32–3, 36, 70, 73, 142 fn
Bucklin system 203–206
cardinal-weighted pairwise comparisons 235–6
choice consistency
  as a condition for RBVPRs 149, 151
  in the Arrow theorem 126–7, 137–8, 140
Clarke, Edward xiv–xv, 292
collective decision
  definition 5–6
  taxonomy of procedures for 9
collectivity, definition 4–5
combinativity 158, 165–242
Condorcet, Marquis de
  development of Condorcet rule 182
  justification for majority rule 84
  supporter of Condorcet consistency 195
Condorcet consistency 154, 158–9, 163, 165–242
Condorcet rule 102–103, 106, 182–90
consensus 10–11, 52–3, 64
contest as a mode 17–18, 55, 68
Coombs, Clyde 213
Coombs rule 213–15
Copeland rule 206–209
cost of computation
  definition 161
  in rules for continuums of two or more dimensions 247–61
  in rules for discrete options 165–242
criteria for evaluating collective decision procedures
  efficiency 36–7
  equity 37–9
  stability 39–40
cumulative voting 265
de la Vallee Poussin, D. 316
decision
  compared to judgment 6–7
  definition 3–4
demand-revealing process
  multiple discrete options 304–10
  one-dimensional continuums 311–16
  two options 292–302
Dodgson, Charles 195
Dodgson rule 196–9
dominance 153
Dreze, J. 316
Droop, H.R. 270–72
dueling 17–18
Dummett, Michael 268–9, 278–9

ease of use
  definition 161
  in rules for continuums of two or more dimensions 247–61
  in rules for discrete options 165–242
estimated centrality rule 215–17
extortion as a mode 12–13, 54, 66
Gibbard–Satterthwaite theorem 143–8
GOCHA set 154–5
Green, Jerry 144
Green-Armytage, James 235–6
Gregory, J.B. 273
Hare, Thomas 270
Hicks, John 28–31
Hill, David 100, 274, 278
Hill, Roland 269
Hill, Thomas Wright 269–70
homogeneity 157–8, 165–242
independence of clones 159–60, 165–242
independence of irrelevant alternatives
  as a condition for RBVPRs 149–51
  in the Arrow theorem 132–3
invariant loss consistency 155, 165–242
Kaldor, Nicolas 28–31
Kemeny rule 182
knowing
  applied to collective decisions 62–8
  definition 57–61
Laffont, Jean-Jacques 144
Larson, J.A.O. 83–4
lucidity 161, 165–242
majority consistency 153, 165–242
majority rule
  justifications for 84–91
    equal weights 84–8
    unequal weights 88–91
    disinterested judges 88–90
    self-interested advocates 90–91
Malinvaud, E. 317
Mann, Irwin 272
Margolis, Howard 294, 316
material equality 49–50
maximin rule 212–13
May, K.O. 84–6
median voter rule
  in multiple dimensions 248–51
  in one dimension 245–7
Meek, Brian 274–5
minimum sum of distances rule 254–6
minimum sum of forces rule
  multi-stage 259–60
  single stage 256–8
monotonicity
  as a condition for RBVPRs 149–50, 157, 162, 165–242
  connection to Arrow's positive association 132
  in majority rule 87
multidimensional balanced pulls rule
  multi-stage 258–9
  single-stage 251–4
multidimensional continuums 247–61
multiple vote plurality rule 264–5
mutual majority consistency 153, 165–242
Nanson, E.J. 195, 201
Nanson rule 201–203
neutrality
  as a condition for RBVPRs 149–50
  in majority rule 85
non-dictatorship
  as a condition for RBVPRs
  in Arrow theorem 133, 139–40
  in Gibbard–Satterthwaite theorem 143, 148
one-dimensional continuums
  demand-revealing process 311–16
  rules with equally weighted votes 245–7
  Thompson insurance mechanism 311
outcome efficiency 46–8
Pareto, Vilfredo 24–6
Pareto consistency
  as condition for RBVPRs 149–50
  in Arrow theorem 134, 140
  in Gibbard–Satterthwaite theorem 146, 148
  relation to *k*-PSC in STV 269
Pareto improvement
  as criterion for collective decisions 24–6

Pareto improvement (continued)
relation to criterion of stability 39
relation to demand-revealing process 297–9, 310
party-list systems 265–7
perfect-tie responsiveness 157, 165–242
plurality 165–70
Polinsky, Mitchell 31
positive responsiveness
as condition for RBVPRs 149, 151, 156, 158, 162–3, 165–242
in majority rule 85–7
preservation 51–2
procedural efficiency 43–6
productivity-based inequality 50
properties of vote-processing rules, table 236–7
proportionality for solid coalitions 268–9
pseudo-consensus 11–12, 53, 64–5
random process 16–17, 54, 67
range voting 174–6
ranked pairs rule 219– 23
Rawls, John 32–3
RBVPR, defined 137
reciprocity 21
resolvability
as a condition for RBVPRs 149, 151, 156–7, 162, 165–242
in majority rule 87
resistance to strategy
definition 161–2
in rules for continuums of two or more dimensions 247–61
in rules for discrete options 165–242
restoration 50–51
Samuelson, Paul 26–8
Schulze, Markus 228, 230
Schulze method 228–32
Schwartz consistency 154–5, 165–242
Scitovsky, Tibor 30
Senatorial rules 273
simplified Dodgson rule 199–201
single transferable vote 267–79
comparison of versions 276–9
paired comparisons by 279–86
Smith consistency 154, 165–242
social welfare function 27–8
stability 39–40
strategy proofness
as a condition for RBVPRs 149, 151
in the demand-revealing process 292–3
in the Gibbard–Satterthwaite theorem 143–8
supportable vote-processing rules 236, 238–42
symbolic equality 48–9
Thompson, Earl 290–91
Thompson insurance mechanism
multiple discrete options 302–304
one-dimensional continuum 311
two options 291–2
trade as a mode 12, 53–4, 65–6
Tullock, Gordon 32–3, 36, 263, 292
two- or more-dimensional continuums
single-stage rules for 247–58
multi-stage rules for 258–60
two-ballot majority rule 190–93
universal domain
as a condition for RBVPRs 149, 151–3, 162, 165–242
in majority rule 85
Vickrey, William 143
voting
attraction of, in light of the Arrow theorem 140–1
compared to other modes 54–5, 67–8
definition of 17
dimensions of 78–80
elements of 77–78
historical origin of 83–84
voting cycles 93–122
conditions that preclude and promote 95–9
frequency of 99–115
probability tables for 116–122
significance of 115
Warren, C.H.E. 274–5
weighted Condorcet rule 223–7
Young rule 217–18
Zeckhauser, Richard 163 fn, 260